

**The
suprasegmental signaling of attitude
in German and Chinese:
A phonetically oriented contribution to intercultural
communication**

**Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Philosophischen Fakultäten
der Universität des Saarlandes**

**vorgelegt von
Patricia E. Mueller-Liu, aus London**

Saarbrücken, 2004

**Band 1/2
Text**

Der Dekan Univ. Prof. Dr. phil. Wolfgang Schweickard

**Berichterstatter: Prof. Dr. William J. Barry
Prof. Dr. Neal R. Norrick
Prof. Dr. Paul M. Thompson**

Tag der letzten Prüfungsleistung: 11.05.2004

Table of Contents

	General Introduction	i
	<i>Part One - Discourse Analysis: The study of conversational interaction</i>	
	Introduction to Part One	2
1	Interactional Sociolinguistics - The study of interpersonal and intercultural communication	6
1.1	John Gumperz: The signaling of conversational involvement and the ambiguity of the contextualization cues	6
1.1.1	Contextualization - the signaling of speaker meaning and intent.....	6
1.1.1.1	The signaling of discourse cohesion - the suprasegmental cohesion markers	7
1.1.1.2	Contextualization - the contextualization cues.....	7
1.1.2	The inferential process - the interpretation of speaker intent	9
1.1.3	The role of linguistic and socio-cultural background knowledge	10
1.1.4	Intercultural communication	11
1.1.5	Language and social identity	12
1.2	Deborah Tannen: The interaction strategies and the ambiguity of the metamessage	13
1.2.1	The metamessage of speech	14
1.2.1.1	Involvement and independence	14
1.2.1.2	Footing - Power and solidarity	15
1.2.2	Interaction strategies	16
1.2.2.1	Brown and Levinson (1978/1987) - Positive and negative politeness.....	16
1.2.2.2	Lakoff's Three Rules of Politeness (1973)	17
1.2.2.3	A fourth interaction strategy: <i>Support</i>	17
1.2.3	The interaction strategies in interpersonal and intercultural communication	18
1.2.4	The use of suprasegmental features in the signaling of interaction strategies	22
1.2.5	The ambiguity of the metamessage	23
1.2.6	The origins of the ambiguity of the metamessage.....	24
1.2.6.1	The sociolinguistic origins - power and solidarity.....	24
1.2.6.2	The linguistic origins - the ambiguity of the suprasegmental features.....	25
	Summary of Part One	26
	<i>Part Two - Dimensions of meaning in suprasegmental communication</i>	
	Introduction to Part Two	29
2	Forms and functions of English and German intonation	33
2.1	The communicative forms of English and German intonation	33
2.1.1	Units of tonality - the intonation-group	33
2.1.2	The onset, the nucleus and the intonation-group boundary.....	34

2.1.2.1	The onset	34
2.1.2.2	The nucleus	35
2.1.2.3	The intonation-group boundary	35
2.1.3	The pitch-patterns of the prenuclear and nuclear segment	36
2.1.3.1	The prenuclear pitch-patterns (<i>heads</i>)	36
2.1.3.2	The nuclear pitch-patterns	37
2.2	The communicative functions of English and German intonation	39
2.2.1	The functions of tonality and tonicity	39
2.2.1.1	Discourse cohesion - the informational functions of tonality and tonicity..	39
2.2.1.2	The grammatical functions of tonality and tonicity	40
2.2.1.3	The para- and extralinguistic functions of tonality and tonicity	41
2.2.2	The functions of nuclear direction.....	42
2.2.2.1	Discourse cohesion - the linguistic functions of nuclear direction	42
2.2.2.2	The paralinguistic functions of nuclear direction	42
2.2.2.3	The extralinguistic functions of nuclear direction	45
2.2.3	The communicative functions of nuclear pitch-range	47
2.2.3.1	Types of nuclear pitch-pattern in English and German	47
2.2.3.2	The paralinguistic functions of the nuclear pitch-patterns.....	48
2.2.3.3	The extralinguistic functions of the nuclear pitch-patterns.....	50
2.2.4	The communicative functions of tail-slope configurations	51
2.2.5	The functions of the prenuclear pitch-patterns.....	51
2.2.5.1	The paralinguistic functions of the prenuclear pitch-patterns	52
2.2.5.2	The extralinguistic functions of the prenuclear pitch-patterns	53
3	The communicative forms of Mandarin Chinese intonation	54
3.1	The Mandarin Chinese tones	54
3.1.1	The lexical tones in isolation - basic configurations	54
3.1.2	The tones in tonal compounds.....	56
3.1.2.1	The rules of prominence	56
3.1.2.2	The rules of tone <i>sandhi</i>	57
3.2	Tonality in Mandarin Chinese	58
3.2.1	The Mandarin Chinese intonation-group	58
3.2.2	The signaling of intonation-group, tone-group and tone-unit boundaries ..	59
3.3	Tonicity in Mandarin Chinese	64
3.4	Intonational pitch-patterns in Mandarin Chinese	66
3.4.1	Utterance-final intonational pitch-patterns on sentence-final particles.....	66
3.4.2	Utterance-final intonational pitch-patterns without sentence particles.....	68
3.4.3	Intonational pitch-patterns across Mandarin Chinese utterances.....	69
3.5	Summary: The communicative forms of Mandarin Chinese intonation.....	72
4	The communicative functions of overall pitch, loudness, tempo and fluency	73
4.1	Definitions - Overall pitch, loudness, tempo and fluency.....	73
4.1.1	Overall pitch	73
4.1.2	Overall loudness	73
4.1.3	Overall tempo	73
4.1.4	Fluency	74
	The communicative functions of overall pitch, loudness, tempo and fluency	
4.2	fluency	75
4.2.1	The linguistic functions - discourse cohesion	75

4.2.2	The paralinguistic functions	75
4.2.3	The extralinguistic functions of overall pitch, loudness, tempo and fluency	76
4.2.3.1	Pitch and pitch-variability	76
4.2.3.2	Overall loudness	77
4.2.3.3	Speech tempo and fluency (pause phenomena)	78
5	Forms and functions in voice quality communication	79
5.1	The communicative forms of voice quality - the articulatory settings	79
5.1.1	The neutral setting	79
5.1.2	The supralaryngeal settings	80
5.1.2.1	The longitudinal settings	80
5.1.2.2	The latitudinal settings	81
5.1.2.3	The velopharyngeal settings	81
5.1.3	The phonatory settings	82
5.1.3.1	Falsetto	82
5.1.3.2	Whisper	82
5.1.3.3	Creak	82
5.1.3.4	Harshness	83
5.1.3.5	Breathiness	83
5.1.4	The tension settings	84
5.1.4.1	Tense voice	84
5.1.4.2	Lax voice	84
5.2	The communicative functions of voice quality	87
5.2.1	The paralinguistic functions of voice quality	87
5.2.1.1	The paralinguistic functions of the phonatory settings	87
5.2.1.2	The paralinguistic functions of the supralaryngeal settings	89
5.2.2	The extralinguistic functions of voice quality	89
5.2.2.1	The extralinguistic functions of phonatory voice quality	89
5.2.2.2	The extralinguistic functions of the supralaryngeal settings	90
	Summary and conclusions of Part Two	91

Part Three - Experimental work: The signaling of interpersonal attitude in German and Chinese

	Introduction to Part Three	94
6	The interactional macroanalysis	97
6.1	The dialogs - a broad outline	97
6.1.1	The recordings	97
6.1.2	The participants	98
6.1.3	The assignment	98
6.1.4	The negotiation of the interactions	99
6.1.5	The interactional events	100
6.1.5.1	Asking for information	100
6.1.5.2	Special wishes and offers	101
6.1.5.3	Refusals and offers of substitutes	101
6.1.5.4	Discussions	101
6.1.5.5	Criticism of the merchandise	102
6.2	The negotiation of the interactional events	103
6.2.1	The interactions with familiar women	103

6.2.2	The interactions with unknown women	106
6.2.3	The interactions with familiar men	110
6.2.4	The interactions with unknown men	112
6.2.5	Summary and conclusions.....	116
7	The sociolinguistic judgement tests	119
7.1	The general procedure of the judgement tests	119
7.1.1	The participants	119
7.1.2	The speech material.....	119
7.1.3	The procedure of the tests	120
7.1.4	The questionnaires	120
7.2	The evaluation of the judgement tests - the quantitative analysis.....	121
7.2.1	The classification of the speech samples.....	122
7.2.1.1	Procedure.....	122
7.2.1.2	The mean scores for footing, valence and strength of position.....	122
7.2.1.3	The ambiguity of the speech samples.....	123
7.2.1.4	The classification of the speech samples.....	127
7.2.2	The influence of the discourse context.....	127
7.2.2.1	The influence of discourse context on judgements of <i>footing</i>	127
7.2.2.2	The influence of discourse context on judgements of <i>valence</i> and strength of position.....	129
7.2.2.3	The influence of the discourse context on standard deviation.....	131
7.2.2.4	Summary: the influence of the discourse context.....	132
7.2.3	Summary and conclusions.....	133
8	The phonetic microanalysis	135
8.1	The analytic listening tests	135
8.1.1	General procedure.....	135
8.1.2	The participants.....	135
8.1.3	The speech material and the questionnaires.....	136
8.1.4	The results of the listening tests.....	137
8.2	The instrumental speech analyses.....	137
8.2.1	Procedure and methodology of the instrumental analyses.....	137
8.2.1.1	The instrumental analysis of phonatory voice quality	138
8.2.1.2	The instrumental analysis of the overall features.....	138
8.2.1.3	The instrumental analysis of intonation.....	139
8.2.1.4	The instrumental analysis of supralaryngeal voice quality.....	141
8.3	The final results of the phonetic microanalyses.....	141
8.3.1	The final results for phonatory voice quality	141
8.3.2	The final results for supralaryngeal voice quality.....	143
8.3.3	The final results for overall pitch, tempo and loudness.....	143
8.3.4	The final results for fluency/non-fluency.....	144
8.3.4.1	The final results for normal and overly long (non-) phonological pause phenomena.....	145
8.3.4.2	The results for speech errors.....	146
8.3.4.3	The final results for overly short and absent phonological pause phenomena.....	146
8.3.4.4	The final results for utterance-final slow-down.....	146
8.3.4.5	The final results for fluency and non-fluency.....	147
8.3.5	The final results for intonation.....	151

9	The phonetic exponency of the interaction strategies.....	166
9.1	The signaling of <i>deference</i> , <i>arrogance</i> and <i>support</i> in the German speech samples.....	166
9.1.1	The signaling of <i>deference</i> , <i>arrogance</i> and <i>support</i> in Clara's speech samples.....	166
9.1.2	The signaling of <i>deference</i> , <i>arrogance</i> and <i>support</i> in Sybille's speech samples.....	168
9.1.3	<i>Deference</i> , <i>arrogance</i> and <i>support</i> in German.....	169
9.1.3.1	The <i>deference</i> strategy in German.....	169
9.1.3.2	The <i>arrogance</i> strategy in German.....	169
9.1.3.3	The <i>support</i> strategy in German.....	170
9.1.4	The role of the suprasegmental phenomena in German.....	171
9.1.4.1	The role of phonatory voice quality.....	171
9.1.4.2	The role of overall loudness, tempo and fluency.....	173
9.1.4.3	The role of overall pitch.....	174
9.1.4.4	The role of supralaryngeal voice quality.....	175
9.1.4.5	The role of intonation in German.....	175
9.1.5	Summary: The phonetic exponency of <i>deference</i> , <i>arrogance</i> and <i>support</i> in German.....	180
9.1.6	The German utterances signaling ambiguous <i>valence</i>	181
9.2	The signaling of <i>deference</i> , <i>arrogance</i> and <i>support</i> in the Chinese speech samples.....	182
9.2.1	The signaling of <i>deference</i> , <i>arrogance</i> and <i>support</i> in Wu's speech samples.....	182
9.2.2	The signaling of <i>deference</i> , <i>arrogance</i> and <i>support</i> in You's speech samples.....	183
9.2.3	<i>Deference</i> , <i>arrogance</i> and <i>support</i> in Chinese.....	184
9.2.3.1	The <i>deference</i> strategy in Chinese.....	184
9.2.3.2	The <i>arrogance</i> strategy in Chinese.....	185
9.2.3.3	The <i>support</i> strategy in Chinese.....	185
9.2.4	The role of the suprasegmental phenomena in Chinese.....	186
9.2.4.1	The role of phonatory voice quality.....	186
9.2.4.2	The role of overall tempo, loudness and fluency.....	187
9.2.4.3	The role of overall pitch.....	189
9.2.4.4	The role of intonation in Chinese.....	189
9.2.4.5	The role of supralaryngeal voice quality.....	195
9.2.5	Summary: The phonetic exponency of <i>deference</i> , <i>arrogance</i> and <i>support</i> in Chinese.....	196
9.2.6	The Chinese utterances signaling ambiguous <i>valence</i>	197
9.3	Comparison: The signaling of speaker attitude in German and Chinese....	198
9.3.1	The signaling of <i>deference</i> , <i>support</i> , <i>arrogance</i> and <i>false deference</i> in German and Chinese.....	198
9.3.2	The role of the suprasegmental features in German and Chinese.....	200
9.3.2.1	The role of phonatory voice quality in German and Chinese.....	200
9.3.2.2	The role of overall tempo, loudness and fluency in German and Chinese..	200
9.3.2.3	The role of overall pitch in German and Chinese.....	200
9.3.2.4	The role of supralaryngeal voice quality in German and Chinese.....	200
9.3.2.5	The role of intonation in German.....	201
9.3.2.6	The role of intonation in Chinese.....	201
9.3.3	Final summary and conclusions.....	201

10	Final conclusion: The signaling of interpersonal attitude in German and Chinese	208
10.1	Interaction strategies in German and Chinese conversational interaction...	208
10.2	The signaling of interpersonal attitude in German and Chinese.....	209
10.3	Implications of the obtained results for intercultural communication.....	211
	Bibliography.....	213
	Curriculum Vitae.....	228

General Introduction

This work is a comparative examination of the signaling of interpersonal attitude by the suprasegmental features¹ in German and Mandarin Chinese. As such it brings together, via a phonetic interface, two disciplines in general linguistics - sinology and interactional sociolinguistics - which have had only limited contact in the past. In order to smooth the path towards this new relationship, this introduction begins with an outline of the most important concepts and achievements of anthropologically-oriented sociolinguistics and, in particular, the research conducted in intercultural communication, a field of interest to sociolinguistics and sinologists alike. Special emphasis is placed on the relevance of phonetics/phonology and the manner in which this discipline can contribute to the solution of the problems encountered in intercultural communication.

Intercultural communication between China and the West

Thanks to the achievements of modern technology, the world is getting smaller by the day. Journeys which took weeks by ship in the early 20th century are covered by plane in a matter of hours today. Faraway places unreachable thirty years ago to most people are now popular holiday resorts. And with the internet and electronic mail, messages to the remotest corners of the earth, often inaccessible by normal postal services, arrive in their receiver's mail box within a few seconds. On the other hand, not only are distances becoming shorter, people are also moving closer to each other; With the dismantling of trade barriers, the establishment of free trade areas and the introduction of common currencies companies from all corners of the globe are in business with each another. And as people nowadays tend to follow their work wherever it may lead them, the globalization of the workplace is leading to an increasing internationalization of society. As a result, in today's global village the representatives of British, German, French and American businesses are quite as much part of city life in Peking, Hongkong and Shanghai as the employees of Chinese, Pakistani or Indian businesses in London, Paris, San Francisco and New York. Even in Germany, which for a long time was always a few steps behind other countries with respect to internationalization, one now frequently encounters Turkish, Indian and Chinese employees, even in domestic companies, as a result of recent economic policies.

In such a small world, where people from so many different countries live and work side-by-side, one might assume that human communication is thriving. Unfortunately, this is not the case. While contacts are established with swiftness and ease, communication between the members of different countries and cultural spheres has not made much progress since the early 19th century when the first western treaty ports were being established across Southeast Asia. Whatever the level of contact involved, the American CEO in Shanghai, the Russian émigré in London and the Indian computer programmer in Munich all grapple with the same difficulties each day - the problems of intercultural communication, involving people speaking what appears to be the same language, but whose use of this language differs in a number of important aspects due to their different socio-cultural origins. While miscommunications encountered in intracultural² communication generally involve simple lexical mismatches which can be dealt with swiftly, the problems of intercultural communication originate in the fundamental incompatibility of basic communicative norms operating below the normal level of consciousness. Therefore, the processes which lead to miscommunication here often go unnoticed by conversationalists and also do not normally diminish with increasing contact but tend to increase over time, developing into interpersonal problems and cross-cultural stereotyping.

¹ In this work, the American term *suprasegmental* is preferred to the term *prosodic*, in general use in British linguistics.

² The term *intracultural communication* relates to communication between speakers of the same language, or variant, adhering to the same interactional and communicative norms.

As globalization increases, the problems of intercultural communication are becoming more widespread and more noticeable and the last decade has seen a strong increase in scholarly writings on topics such as intercultural management, cross-cultural business communication and intercultural communication at the workplace (such as Bergemann et al.1992, Mead 1993, Scollon 1995, Clyne 1996 and Gudykunst et al.1996). On a parallel plane, linguists, in particular those working within the anthropologically oriented disciplines, are becoming interested in the linguistic problems of intercultural communication - those involved in the actual communication process. Though interest in the interface between language and culture goes back a long way, beginning with the work of the American linguists Franz Boas and Edward Sapir in the early 19th century³, the strength of the bond between language and culture was not realized before the second half of the last century, when the systematic research of these issues began. The research conducted within the ethnography of communication and interactional sociolinguistics under the direction of Dell Hymes and John Gumperz revealed the degree to which language, culture and society are interconnected and the extent to which the language of an individual is determined by the socio-cultural norms of his speech community. Thus, for example, Gumperz' pioneering studies of conversations between speakers of Indian and western (British) English, and speakers of Black English Vernacular and Standard American (1972, 1972a, 1977, 1982a, b) showed that in spite of what was thought to be a 'common language', in these cases British and American English, serious miscommunication nevertheless occurred because the communicative norms and conventions which regulated the conversationalists' use of this language were not shared due to their different socio-cultural origins.

Gumperz' early findings have been replicated by other studies on intercultural communication conducted by Gumperz and his associates, such as Kochman 1981, on communication between speakers of Black English Vernacular and Standard American, Scollon 1985, 1995 on conversational interaction in Athabaskan and Lapp communities and Tannen 1984, 1986⁴, on communication between Americans of different ethno-cultural extraction. While differences in interaction, or politeness, norms are also to blame, the results of these and related studies indicate that the greatest problems are in fact created by the suprasegmental features of speech⁵, involving the intonation, loudness, pitch and tempo of utterances and the speaker's voice quality⁶. It was found, for example, that speakers automatically transfer the subconscious rules regulating the use of the suprasegmental features in their mother tongue to other languages, this leading to problems in both signaling and interpretation processes. A second, perhaps even more important factor at the origin of communication problems involving these features is their functional diversity, by virtue of which, as Trubetzkoy has observed, "When we hear someone speak, we not only hear what he says, but can also tell what mood he is in and who he is"(1939:41)⁷. This means that the suprasegmental features can serve to signal at least three different kinds of communicative meaning: (1) *linguistic* meaning, communicating information which relates to the structure of the speaker's message; (2) *paralinguistic* meaning, relating to his emotional and attitudinal disposition⁸; and (3) *extralinguistic*, or

³ Sapir's observations, based on evidence from studies of American Indian languages, led to the Sapir-Whorf hypothesis which holds that a language determines the way its speakers conceptualize and structure the world around them. See Fasold 1990 and Anderson 1985.

⁴ Linguists in Germany working in this paradigm are Couper-Kuhlen (1988, 1993, 1996), Auer (1986) and Hinnenkamp (1989, 1997).

⁵ In this work all these suprasegmental features are collectively referred to as 'suprasegmental features'. The frequently encountered distinction between 'prosodic' and 'paralinguistic' features, the former used to refer to communicative phenomena based on variations in *length*, *loudness*, *pitch* and *pause*, the latter denoting features of voice quality, is not applied here. (See Crystal 1969,1975).

⁶ The phenomenon of voice quality is also referred to in the literature as the speaker's *tone of voice*.

⁷ The German original is: 'Wenn wir jemanden reden hören, so hören wir *wer* spricht, in *welchem Tone* er spricht und *was* er sagt'.

Unless otherwise indicated, translations into English are by the writer.

⁸ The term *emotion* is used to refer to internal dispositions, or *feelings*, of an individual, such as anger, joy and fear, generally agreed to be externalized by unmonitored, non-cognitive means of expression originating in universal physiological processes. (See Crystal 1969, 1975, Scherer 1979a, 1995) The term *attitude*, by contrast, is seen as a position, or stance, assumed with respect to an event, message or person, and signaled by cognitively monitored, conventionalized modes of expression which are largely language-dependent (Scherer 1979a, Couper-Kuhlen 1986).

indexical, information, relating to long-term qualities of the speaker, such as his socio-cultural origins, his level of education and profession, and certain traits of his personality⁹. Though at a first glance the functional *diversity* of the suprasegmental features might seem conducive to human communication, since it means that a small number of communicative signals may be used to communicate a large range of different meanings, in effect, however, this diversity poses enormous problems, as it also means that one and the same signal can simultaneously communicate a very large range of different meanings. This functional *ambiguity* is aggravated by the fact that, as Tannen has remarked, these signals “do not come with instructions for use” (1986: 85), so that the precise meaning of a suprasegmental feature - as intended the speaker - is often misinterpreted by the listener, unable as he is to peer into the speaker’s mind and see how it was intended. While this factor already suffices to cause misunderstandings in intracultural communication, involving conversationalists speaking the same language, in intercultural encounters the problem is aggravated by the fact that the socio-culturally patterned rules which regulate the use of these features often differ strongly from one speech community to the next. As a result, the interpretation of speaker meaning in intercultural suprasegmental communication often resembles a guessing-game.

In comparison to the large body of research available on intercultural communication between speakers of different western languages, intercultural studies on the languages of Africa, Asia and the Middle East are scarce. As a case in point, despite the rapidly growing importance of China on the economic and political stage, interculturally-oriented linguistic studies on the languages of China are virtually non-existent. While this state of affairs is also due to the relative isolation of the Peoples’ Republic, a second, probably more important reason lies in the Chinese language which, due to its complexity, remains inaccessible to most workers in intercultural communication. As a third factor, the research literature on Chinese shows a strong and widespread emphasis on matters of grammar and phonology, coupled with what seems to be an aversion to the suprasegmental features, so that only very little is known on suprasegmental communication in Chinese. As a result, the few cross-cultural studies that have been conducted on Chinese have had to content themselves with the analysis of conversations held in western languages, such as English (Young 1982, 1996; Scollon 1995) and German (Günthner 1993), and discussions have been widely limited to communication difficulties relating to pronunciation and grammar. On the other hand, a few studies, such as Young (1982) and Günthner (1993)¹⁰, have touched upon problems potentially involving the suprasegmental domain. The picture which emerges is that the problems encountered here are quite as frequent, serious and far-reaching as those which have been found to occur in encounters between speakers of Indian and western (British) English.

As Chinese is a tone-language, in which pitch is used not only for suprasegmental communication but also, in the form of the lexical tones, to signal different lexical items, misunderstandings of pitch in this language start on the tonological domain, involving the phonological use of pitch in the distinction between different words¹¹. As an illustration of this kind of misunderstanding, examples of which are often described in the literature, a young British missionary in Wenzhou in 1924 was surprised at first that the native priest did not seem to understand her enquiry after the health of his wife who had recently fallen ill - until it turned out that, on account of an error in lexical tone, she had in fact asked about the recovery not of his spouse, but of his cow¹². While such misunderstandings already represent an impediment

⁹ These definitions of the terms *linguistic*, *paralinguistic* and *extralinguistic* are adopted from Laver 1980 and Trench 1990.

¹⁰ Young (1982) and Günthner (1993) focus on intercultural differences with respect to what Gumperz refers to as ‘discourse strategies’ between Chinese and English and German, respectively. These relate to the different ways messages are structured, information is presented and topics are dealt with in conversation. Günthner also includes a discussion of the Chinese ‘conversational style’ or *habitus*, including some of the many proverbs and idiomatic expressions Chinese tend to weave into their speech. The terms *conversational style* and *habitus* are explained in Chapter 1.

¹¹ The lexical tones of Mandarin Chinese and their phonological functions are discussed in Chapter 3.

¹² This misunderstanding occurred in the southwestern Chinese dialect of Wenzhou. Unless indicated otherwise, examples given in this work describe genuine events and relate either to personal experiences or observations of the writer, or events recounted to her by friends and acquaintances.

to smooth communication between China and the West, misunderstandings on the suprasegmental level, involving the suprasegmental use of pitch, highly difficult to detect and therefore only rarely mentioned in the literature, can have truly disastrous consequences. Thus, a Chinese man married to a German woman was never able to understand why, when they were having a discussion or she was explaining something to him, she frequently spoke in a high-pitched, loud and seemingly irritated voice. Since the use of the suprasegmental features did not form part of the curriculum of the Chinese courses the young woman had attended at university, she did not realize until years later, when it was much too late to make repairs, that in contrast to most western languages, where increases in pitch and loudness are employed to signal emphasis, in Chinese these features are used exclusively to communicate states of heightened emotional arousal, as in irritation and anger¹³. As the origin of the problem could not be found - the young man was never able to explain what exactly was bothering him - this mismatch between the Chinese husband's and the German wife's use of the suprasegmental features ended up being one of the factors leading to their separation. These examples demonstrate that the problems encountered in intercultural misunderstandings involve not only misinterpretations of referential meaning, which are comparatively easy to repair, but also, far more seriously, misinterpretations of the speaker's interpersonal attitude, signaling his position with respect to the listener and the degree to which he is able to identify himself with his qualities, wants and needs¹⁴. Thus, the increased pitch and loudness of the German wife's voice when she was explaining something to her husband meant to him - i.e. according to his communication system - that she was annoyed with him. Given the serious consequences of such misunderstandings of attitude, miscommunication originating in the suprasegmental domain is of great concern to all workers in interactional sociolinguistics.

In contrast to the communication of emotion, which is achieved by signaling processes originating in universal physiological processes and does not normally involve misunderstandings of the kind discussed here, the communication of attitude is primarily associated with intonation, whose communicative forms and functions are language-dependent and often vary from one language to the next (Fónagy & Bérard 1972, Ladd et al. 1985, Bolinger 1986)¹⁵. The importance of intonation in the communication of interpersonal attitude can also serve to explain the scarcity of research in intercultural communication involving the Chinese language. In contrast to the view of some early scholars, such as Karlgren (1918/75), it is now widely accepted that Chinese does possess something which may be referred to as sentence intonation, first sketches of which are found in Chao 1968 and Kratochvil 1968. However, researchers are still mostly grappling with the problems of *tonation*, the interplay between the lexical tones, so that even the most basic issues of Chinese intonation, such as the signaling of the focus, the syntactic correlates of the intonation-group and the role of declination in Chinese have not been clarified to any degree of satisfaction. Given the slow progress in the study of Chinese intonation, we are no nearer today to answering even the most basic questions on the nature of Chinese intonation than forty years ago when intonation was declared to be universal, but nothing precise was known about its functionings in Chinese (Bolinger 1964, 1978)¹⁶. Naturally, as long as the most basic questions on the forms and

¹³ Naturally, misunderstandings involving the semantic and pragmatic use of words and expressions are also frequent between China and the West. Thus, on a different occasion this same German wife offered to lend her Chinese husband an English-Chinese dictionary. Instead of being pleased, however, he got very angry at this offer, a fact which made her feel very unhappy and insecure - until he explained that in the traditional Chinese view, husbands and wives belong together body and soul and do not use the expression 'lend' with each other.

¹⁴ This definition of interpersonal attitude is based on Deborah Tannen's concept of interpersonal attitude, which she refers to as the *meta-message of speech* (1984, 1986). See Chapter 1.

¹⁵ Intonation is defined as relating to cognitively-monitored variations in pitch extending over whole utterances and clearly detectable by the human ear. These pitch-variations form, or *trace*, the intonation-contours of utterances, to be discussed in detail in Chapter 1. Smaller-scale variations in pitch limited to individual segments and scarcely discernable by the human perceptory system are referred to as (pitch-) *jitter*. In contrast to jitter and overall pitch (see Chapter 4), it is generally agreed that the cognitively-monitored pitch-patterns of intonation do not directly participate in the expression of emotions (Bolinger 1986, Mozziconazzi 1998).

¹⁶ As one example of early research in intonation, Bolinger's (1978) discussion of intonation across languages includes only very brief mention of Chinese, in which it is merely stated that in the (non-Mandarin) dialect of Lungtu the pitch-patterns of utterance-final tones are 'affected by intonation'. (1978: 496). No sources are given for this information, however. How little is known on suprasegmental commu-

functions of intonation in Chinese have not been answered, the signaling of interpersonal attitude in this language cannot be studied. In order to make such research possible in the future, part of the present work is devoted to developing a functionally-oriented framework of Mandarin Chinese intonation, based on the outlines provided by Chao (1968) and Kratochvil (1968, 1998)¹⁷.

Organization of the book

The present work falls into three parts, the first two of which constitute the - sociolinguistic and phonetic/phonological - framework of the experimental study presented in Part Three. Part One begins with a brief introduction to *discourse analysis*, the most important working methodology applied today in the study of interpersonal interaction. Following this, Chapter 1 begins with an outline of John Gumperz' approach to interpersonal and intercultural communication and introduces the reader to the kind of work conducted in interactional sociolinguistics. In the second half of Chapter 1 we turn to Tannen's work on interpersonal interaction, in particular her model of interpersonal attitude, to be adopted in principle here, but expanded somewhat for the purpose of the experimental study. The chapter concludes with a discussion of Gumperz' and Tannen's findings on the role of the suprasegmental features in the signaling of interpersonal attitude, focusing on the factors which in their view give rise to the ambiguities in the discussed signaling processes.

Part Two (Chapters 2 - 5), representing the phonetic/phonological framework of the present work, introduces the reader to the manner in which the suprasegmental features - intonation, overall pitch, loudness, tempo and fluency, and voice quality - are described and analysed in phonetics/phonology. After a brief general introduction to the forms and functions of suprasegmental communication, each of the three sets of communicative forms is discussed in detail - the forms of intonation in Chapters 2 and 3, the overall forms of loudness, pitch, tempo and fluency in Chapter 4, and the forms of voice quality in Chapter 5. The discussions begin with phonetically oriented definitions of the suprasegmental phenomena under discussion, following which the functions of their communicative forms in German and Chinese are presented on the basis of the research literature. The discussions of the German and Chinese forms of intonation are based on the tonetic approach which places a strong emphasis on the communicative and, in particular, the attitudinal functions of intonation. The tonetic model also has the advantage of being applicable to a wide range of languages, not only English, German and other western languages, but also Mandarin Chinese, to which it is applied in Chapter 3. The discussion of voice quality follows the model set up by John Laver (1980).

The experimental study conducted within the framework of this work is presented in Part Three (Chapters 6 - 10). Combining the methodology followed in interactional sociolinguistics with the analytic methods of experimental and instrumental phonetics, the study involved the recording and analysis of German and Chinese dialogs conducted under naturalistic circumstances, in which different speaker attitudes were elicited¹⁸. After a first sociolinguistically oriented discussion of the recorded dialogs in Chapter 6 (the *macroanalysis*), Chapter 7 describes the sociolinguistic judgement tests which were held with German and Chinese native speakers for a number of utterances taken from the recordings in order to analyse them

nication in Chinese and particularly on intonation - in terms of concrete, systematic communicative patterns applicable in the analysis and instruction of the Chinese language - is illustrated by the fact that the information to be found on this topic in recent grammars of Chinese, such as Norman 1988, Li & Thompson 1989 and Reichhardt 1990, is mostly limited to impressionistic descriptions totaling no more than a few lines. Among earlier works, Chao 1968 includes a detailed and as yet unsurpassed, discussion of a large number of suprasegmental phenomena in Chinese, subsumed under the heading of 'intonation' (1968: 40 ff.). Another example of early work on Chinese intonation is Cheng 1958, on the variant of Mandarin spoken in Szechuan.

¹⁷ Chao's pioneering work on Chinese intonation (1933, 1948, 1968) has unfortunately not been fully appreciated, no doubt due to the strong general emphasis on matters of phonology.

¹⁸ Note that these dialogs are instances of interpersonal, not intercultural, interaction. Linguistic analyses of intercultural communication between German and Chinese interactants are not possible at this point, without knowledge of the suprasegmental processes in Chinese, in particular Chinese intonation.

in terms of the interpersonal attitude signaled in each case. This is followed by the auditory and instrumental analysis (the *microanalysis*) of the suprasegmental features in these utterances, described in Chapter 8 and 9, and in which the phonetic exponency of these different types of attitude - i.e. the manner in which they were signaled by the suprasegmental features - was determined for both languages. The results of the study are summarized in Chapter 10 and their implications for intercultural communication between speakers of German and Chinese are discussed. To the extent to which German (and English) may be regarded as representative for western, i.e. Indoeuropean, languages, the achieved results are also of interest for intercultural communication between China and the West.

Transcription conventions

In the transcribed passages of German and English, stressed and focused/nuclear syllables are highlighted with *cursive* script. The English translations of Chinese passages are in normal, the Chinese originals in *cursive* script, using the *pin-yin* transliteration system, explained in Section 1 of Appendix I. In discussions, figures and tables relating to intonation, stressed and focused/nuclear syllables are differentiated by the additional underlining of focused/nuclear syllables. The tonetic marks used in the representation of German intonation and the symbols used to represent Chinese intonation in Chapters 2 and 3 (and the tables of Appendix III) are presented in the introduction to Section 2 of Appendix III. The transcription conventions used in the other tables of Appendix III are also explained here. In the English and German examples of Chapter 2 minor intonation-group boundaries are represented by a single vertical line (|), major group boundaries by a double line (||). In the Chinese examples discussed in Chapter 3 tone unit boundaries are represented by square brackets ([]), tone-group boundaries by single vertical lines (|), and Chinese intonation-groups by curly braces ({}). The lexical tones are represented with small raised numbers, zero (0) being used for neutral and neutralized tones. *Sandhi*- and stress/prominence-induced changes to tones are separated from original tones by a slash (/).

In Chapter 6, the English (translated) text of German and Chinese passages taken from the dialogs are presented in a manner adapted from the *partitur* - transcription model of Henne & Rehbein 1995, where the text of an utterance, including pauses, is shown towards the left of each line and specifications of loudness, pitch, tempo, fluency, voice quality and the auditory feeling of utterances are given at the right. As in Henne & Rehbein 1995, pauses are indicated using one or more 'plus'- symbols, each standing for a silent period of approximately 0.4 secs. Thus '+' represents the pause of approx. 0.3-0.7 secs., ++ a pause of approx. 0.8-1.2 secs., and +++ stands for a pause of between 1,3 and 1,7 secs. Pauses exceeding this length are given in seconds, placed between square brackets. In the specification of loudness, pitch, tempo and fluency, elevated and reduced values are given in words, mid values not indicated explicitly. Disfluency is indicated with terms such as 'disjointed' and 'hesitant', hyperfluency as 'overly fast' or 'hurried'. In the description of voice quality the following terms are used for the auditory feeling of phonatory *settings*: 'soft' (for *breathiness*), 'warm' and 'gentle' (for *laxness*), 'matter-of-fact' (for modal voice), and 'cool' or 'cold', (for tense *settings*). The auditory effect of *palatalization* is indicated with the terms 'childlike' or 'childish', the presence of smiling (*labiodentalized voice*), detectable auditively by means of its auditory correlates, is indicated as such in words. As these texts are translations of the originals, intonation is not included, but three full stops at the end of utterances (...) represent a 'trailing-off' of pitch, loudness and tempo in the speaker's voice, corresponding to a strong, yet still gradual, utterance-final slow-down, often accompanied by a leveling-off intonation (See Chapters 2 and 9).

Part One

*Discourse Analysis:
The study of conversational interaction*

Introduction to Part One

Discourse analysis, still a relatively new method of linguistic research, has gained great popularity during the last two decades, not only within linguistics proper but also in a large number of its so-called 'hyphenated' subdisciplines, such as sociolinguistics, anthropological linguistics and psycholinguistics. The growing impact of discourse analysis is also felt in other disciplines concerned with human communication in a more general way, as researchers in communication science, artificial intelligence, psychology, sociology and anthropology are also increasingly finding recourse to this methodology in their work. According to Fillmore (1974), this growing interest in the way language is used in social and conversational interaction may be explained by the realization that 'the language of face-to-face conversation is the basic and most primary use of language, all others being best described in terms of their manner of deviation from that base'¹. Among the many disciplines in which discourse analysis is practised today, that of most immediate importance to this work is interactional sociolinguistics. Situated at the intersection between anthropology, sociology and linguistics, interactional sociolinguistics focuses on the concrete processes and events of everyday human interaction, that is, on the failures and successes encountered by conversationalists when using language in interaction (Fasold 1990)². The findings of two further related fields, the ethnography of communication and ethnomethodological conversation analysis, are also of interest to the research conducted within the scope of this work. The present introduction begins with a brief outline of the origins and most important principles of the ethnography of communication and ethnomethodological conversation analysis, after which we come to a more detailed discussion of the theoretical and methodological approach followed in interactional sociolinguistics.

*The Ethnography of Communication*³

The father of the ethnography of communication is the anthropologist Dell Hymes (1962, 1972 a,b, 1974), according to whom an important area of human communication was being omitted in the anthropological and linguistic research of the early 1960s. In the first area, the ethnographically anthropologists were busy studying various aspects of cultural systems, such as kinship systems, religious rituals and healing methods, and tended to regard language as a mere means to achieving their ends. In linguistics, on the other hand, Hymes felt that too much attention was being paid to language as an abstract system of signs, the more concrete - and to him more important - question of how language is used in everyday life, for instance to ask for or communicate information and signal politeness, empathy or disagreement, being almost totally ignored. In Hymes' new discipline therefore, the focus of research shifted to 'the situations and uses, the patterns and functions, of speaking as an activity in its own right' (Hymes 1962: 101), not only with respect to the signaling of referential information and the negotiation of everyday affairs but also more sociologically oriented uses of language, such as the production of social structures and identities. Today, as then, the ethnographer pursues the answers to two basic questions: (1) What does a speaker need to know in order to communicate appropriately within a speech community, and (2) How does he or she learn to do so? (Saville-Troike 1982:2). This knowledge, referred to as *communicative competence*, not only pertains to linguistic rules of communication but also involves a community's norms of social interaction which are widely determined by socio-cultural factors. As the 'ways of speaking', or 'using language', practised by the members of different social and cultural groups often vary considerably, ethnographic research starts with the study of the speech behavior of one

¹ Quoted in Tannen 1984: 6.

² According to Fasold (1990: 65), this interest in concrete interactive events distinguishes this discipline from a second discipline referred to as discourse analysis and practiced, among others, by Van Dijk (1980) and de Beaugrande (1983). Scholars in this field study the structure of - both oral and written - texts. It is therefore also referred to as text linguistics (See Brown & Yule 1983, Stubbs 1980)

³ This discipline is sometimes referred to as the ethnography of speaking. (See Fasold 1990: 40f.)

speech community⁴, following which comparisons with the speech norms of one or more other speech communities are undertaken.

Coming to the methodology followed in this discipline, ethnographic research focuses on natural speech embedded in concrete socio-cultural and communicative contexts. Therefore, ethnographic studies are generally not conducted in speech laboratories but within the natural everyday surroundings of the speech community under investigation. As a second point, since one of the ethnographer's most important aims is to understand the norms and values of the speech community which can account for the (speech-) behavior and attitudes of its members, he must have an intimate understanding of this community. This is achieved by a number of methods, such as *participant observation* and *introspection*, originally adopted from ethnographically oriented anthropology (Fasold 1990). The most important among these is participant observation which involves prolonged stays in the community to be examined, extending from a few months to a year or more. During this time the investigator strives to become part of the community and learns to see the world as the members of this community do (Saville-Troike 1982)⁵. As a third point, ethnographic research is based on three hierarchical units, the *speech situation*, the *speech event* and the *speech act* (Hymes 1972a). The speech situation, not necessarily of a communicative nature, is defined as a situation in human social life associated with the optional presence of speech, such as a healing or wedding ceremony. On the other hand, the speech event, examples of which are a conversation, a lecture or a prayer session, is always communicative and regulated by rules for the use of language. The speech act is defined as the smallest component of a speech event, examples being commands, greetings and summonses (Fasold 1990: 42)⁶. Finally, three further ethnographic concepts commonly encountered in descriptions of speech events are the *setting*, the *scene* and the *participants*. While the setting involves the physical circumstances, such as the time and place, of the speech event, the scene relates to the type of event taking place. The participants include the speaker, the listener and any audience which may be present. Examples of ethnomethodological studies are Ochs' study of a Malagasy village in Madagascar (1973, 1974), Reisman's investigations among Antiguan and Lapps (1974) and Sherzer's study of the Cuna people in East Panama (1974). See also the papers collected in Gumperz & Hymes 1972 and Bauman & Sherzer 1974.

*Ethnomethodological Conversation Analysis*⁷

Ethnomethodological conversation analysis was initiated in the early 1970s by a group of sociologists around Harvey Sacks, Emmanuel Schegloff and Gail Jeffersen, who wished to distance themselves from the 'quantitative' or 'correlative' approach to sociolinguistics which was popular at the time. Instead of studying phonological variation, to be discussed in more detail below, Sacks and his co-workers felt that the object of sociological study should be the communicative means, or *mechanisms*, used by the members of a society to construct and act within their social worlds (Levinson 1983). Research in ethnomethodological conversation analysis examines the mechanisms used by conversationalists to organize and manage the course of conversational interaction. Examples of these are turn-taking mechanisms, and methods of effecting conversational openings, closings and asides, the introduction of topics and the repair of conversational errors and mishaps.

⁴ The speech community is broadly defined as a group of individuals who share the same rules and patterns for 'what to say and when and how to say it' (Fasold 1990: 62). As Saville-Troike (1982) has observed, speech communities can overlap, so that one individual can simultaneously belong to more than one speech community.

⁵ Among the other methods discussed by Saville-Troike are: detached observation, interviewing, interaction analysis and philology, the latter referring to the analysis of written texts (1982: 118 ff.).

⁶ It should perhaps be pointed out that, as Fasold (1990: 42) observes, what is meant by the ethnographic term 'speech act' is not identical to the concept of the 'speech act' used in language philosophy to denote a certain type of communicative intention with a specific effect on the listener (Austin 1962, Searle 1969, 1975. See also Coulthard 1977).

⁷ Detailed discussions of the origins and working principals of ethnomethodological conversation analysis are found in Levinson 1983, Günthner 1993 and Bergman 1981. See also Garfinkel 1967, 1972 and Schiffrin 1989.

With respect to methodology, the ethnomethodologists are strictly empirical. Only naturally occurring conversations are examined, premature theorizing is avoided and analytical categories are developed on the basis of concrete speech data. Given the strong emphasis on the conversational participants' own methods of production and interpretation, which the investigator strives to determine, no methodological procedure has ever been developed, a fact often criticized by researchers in other fields (Fasold 1990, Günthner 1993)⁸. Among the most widely anthologized early research are Schegloff & Sacks 1973, on conversational closings, Sacks, Schegloff & Jefferson 1974, on turn-taking, Jefferson 1972 on side-sequences and Duncan's work on backchannel signals (1973, 1974). Examples of more recent work are Jefferson's work on timing (1973, 1978, 1983) and laughter (1984), Schifffrin's research on discourse markers (1987) and Norrick 1991, on error correction. See also the papers collected in Schenkein 1978, Psathas 1979, Atkinson & Heritage 1984 and Ochs et al. 1996.

Interactional Sociolinguistics

Interactional sociolinguistics or ethnographically oriented discourse analysis (Tannen 1993b), was developed by the anthropological linguist John Gumperz in the 1970s, his conception of this discipline resulting out of earlier ethnographic fieldwork on socio-regional dialects and bilingualism in Europe, India and the United States (1964, 1977, 1982a,b)⁹. In his ethnographically oriented approach to the study of human interaction Gumperz' focus shifts from ritualistic events to everyday conversational interaction. Like the ethnomethodological conversation analysts he examines concrete conversational encounters, but his object of interest, is slightly different. Rather than the means (*mechanisms*) of organizing and managing the course of conversation - strategies relating more to interactive behavior than to the actual use of language - interactional sociolinguistics is interested in the truly linguistic aspects of conversation; Research focuses on the manner in which conversationalists use speech phenomena, including the suprasegmental features, as part of so-called *communicative strategies*, to negotiate conversational interaction and attain their interactive goals (Gumperz 1972)¹⁰. Gumperz' interest in these communicative strategies is twofold: As an anthropologist, his aim is to determine whether and in what way their realization and use differs between different social and cultural groups. On the other hand, from a sociolinguistic perspective, he is interested in the manner in which the communicative context of interactive events and the socio-communicative knowledge of the participants contribute to the signaling and interpretation of speaker meaning and intent¹¹. In interactional sociolinguistics these relate not only to factual, referential, information but also to socially-oriented messages, such as attitudinal information and the signaling of social structures and identities. While the last two points constitute a link to mainstream sociolinguistics, Gumperz, like the ethnomethodologists, distances himself from the methodology applied in the quantitative paradigm, where the use of phonological variables, such as, in English, word-final [-in] in the place of /-ing/ and the use/omission of verb-final [-s] and post-vocalic [-r], by a speaker is quantified and correlated with predetermined social categories like social class, educational status and gender¹². In contrast, Gumperz' interactional approach examines the manner in which speakers actively produce or, in effect, *create*, these social affiliations, or identities, by their use of such variables. As he observes:

⁸ A rough outline of the procedure followed in ethnomethodological studies is provided by Günthner 1993. See also Sacks 1984.

⁹ See Gumperz 1971, 1974, 1977 and Blom & Gumperz 1972.

¹⁰ What Gumperz (1972) refers to as a 'communicative strategy' may be broadly glossed as a socio-culturally patterned manner of using language to achieve a particular communicative goal. In later works, the term he uses more frequently is 'discourse strategy' (1982a,b).

¹¹ Interactional sociolinguistics stresses that harmonious and successful interaction always depends on the cooperation of all participants. (Gumperz 1982a, Tannen 1994; See also Duranti & Brenneis 1986)

¹² Examples of such studies are Labov 1966, 1972, Trudgill 1972, 1974, Milroy 1980 and Cheshire 1982, and the articles collected in Trudgill 1978, Coates 1986 and Holmes 1992. Detailed discussions of this research paradigm are found in Graddol & Swann 1989, Fasold 1990 and Chambers 1995.

[Sociolinguists] customarily take gender, ethnicity and class as given parameters within which we create our own social identities. The study of language as interactional discourse demonstrates that these parameters are not constants that can be taken for granted but are communicatively produced (1982b: 1).

A point of great importance to Gumperz is that the correct interpretation of speaker meaning and intent is only possible if conversationalists share - at least a certain amount of - linguistic and socio-cultural background knowledge, including the rules regulating the communicative use of speech phenomena. This emphasis on the strong link between language use and culture is also found in his approach to the study of intercultural communication, where he examines the manner in which members of different cultures and ethnic groups use language to communicate referential, social and attitudinal information.

The methodological procedure followed in interactional sociolinguistics reflects its strong ethnographic orientation: ethnographic fieldwork, including participant observation, is as much part of the researcher's work as the recording and examination of interactional events. Like the ethnomethodologists, Gumperz works only with naturally occurring speech and does not voice *a priori* statements as to what is likely or expected to occur in interactions¹³. He has also avoided setting up a detailed *modus operandi* of analytical procedure. However, since his reticence here is not quite as strong as that of the ethnomethodologists, a broad outline of the methodological procedure he recommends may be gained from his works. The following is a condensation of what may be found in Gumperz 1982a, supplemented by details provided by his student Deborah Tannen¹⁴.

Step 1: Recording a conversational interaction

A conversational event is tape- or video-recorded with the consent of the participants.

Step 2: Locating passages of interest

The speech material which emerges from the recordings is examined for two kinds of passage of interest: (1) passages in which interaction is clearly smooth, harmonious and pleasing to the interactants, and (2) episodes containing so-called 'uncomfortable moments' potentially reflecting some kind of miscommunication¹⁵.

Step 3: Finding the linguistic and/or socio-cultural origins of the observed phenomena

The investigator searches for culturally patterned differences or common points in the interactants' socio-cultural backgrounds and communicative habits which may be able to account for the observed uncomfortable moments and/or harmonious passages.

*Step 4: Organizing judgements tests for the clarification of the observed phenomena*¹⁶

For the final clarification of the observed communicative phenomena, the passages of interest are played to a panel of judges who share the interactants' socio-cultural background knowledge and communicative habits. Their task is to identify the socio-cultural and communicative processes which gave rise to the observed phenomena¹⁷.

These four steps will be encountered once more in Part Three, where Gumperz' methodology is applied in the first part of the experimental study conducted within the scope of this work. (See Chapters 6 and 7)

¹³ Note, however, that Gumperz does not totally oppose the recording of constructed and loosely controlled encounters. As he writes, 'Where for [certain] reasons, direct recording [of naturally occurring speech] is not possible, actual situations can be recreated through play to gain an insight into the subconscious communicative phenomena [...] if these naturalistic situations are skillfully constructed and not too carefully predetermined' (1982b:11).

¹⁴ See Gumperz 1982a: 30ff. and 132 ff., Tannen 1984: 32 ff., and, in abbreviated form, Tannen 1984: 160f. and 1994: 5 f.

¹⁵ The expression 'judgement test' is proposed here. Gumperz and Tannen have no specific terms for these tests.

¹⁶ As a third type Tannen also names passages that are 'representative of some characteristic that has been noticed' (1984: 160).

¹⁷ According to Tannen (1984), this panel of judges should include the participants in the interaction which is under analysis. This view reflects the ties between interactional sociolinguistics and ethnomethodological conversation analysis.

1. Interactional Sociolinguistics - The study of interpersonal and intercultural communication

The first part of this chapter is devoted to Gumperz's communication theory, including a brief outline of his approach to the study of intercultural communication¹⁸. Section 1.2 discusses Tannen's work on communication, focusing on the signaling of speaker attitude.

1.1 John Gumperz: The signaling of conversational involvement and the ambiguity of the contextualization cues

Gumperz's communication theory deals with the question of how conversational encounters can be made to proceed smoothly and successfully, how to avoid miscommunication and ensure that the interactants' communicative goals are attained. In his view, this is achieved by the establishment of conversational cooperation, which ensures that the conversationalists are 'involved', i.e. know what a given conversation is about, what communicative goals are being pursued and what kind of response is likely to be appropriate to a given speaker message¹⁹. Most importantly, conversational involvement²⁰ enables conversationalists to correctly interpret not only the factual, referential message of an utterance, but also the communicative intent of the speaker, including in particular the underlying social and/or attitudinal message of his words²¹. In order to ensure that the correct interpretation of his communicative intent can take place, the speaker must provide his listener(s) with the information they require, a process referred to as *contextualization*, in which the speaker contextualizes his utterance(s) by signaling their socio-communicative context. While lexical content and syntactic structure also play a certain role in this process, Gumperz takes a particularly strong interest in the contribution of a further group of communicative phenomena, which he refers to as *contextualization cues* (1982a: 131, 1982b: 13). Consisting of code-, style- and dialect-switching processes, formulaic and idiomatic speech and the suprasegmental features²², these *contextualization cues* are special in that they have non-referential, implicit and therefore highly ambiguous meanings. As the use of these phenomena in interpersonal communication often varies between speech communities, Gumperz stresses that in order to interpret correctly the cues he receives, a listener must possess the same - or at least a compatible form of - linguistic and socio-cultural background knowledge as the speaker. Thus, in Gumperz' interaction theory conversational involvement, guaranteeing smooth and successful interaction, relies on three points: (1) contextualization by the speaker, (2) interpretation by the listener, and (3) the common linguistic and socio-cultural background knowledge of the conversationalists²³. In the following we take a closer look at these three phenomena.

1.1.1 Contextualization - the signaling of speaker meaning and intent

The information which the speaker must impart to the listener as part of contextualization is threefold: He must communicate the broader socio-cultural frame of his message, the *speech activity* which he sees himself engaged in, such as, for example, 'discussing politics', 'teaching a French class' or 'shopping at the local grocery store' (1982a: 101 f.)²⁴. Secondly,

¹⁸ The following discussion of Gumperz' interaction theory is based on Gumperz 1982a and b.

¹⁹ The term 'speaker message' is used here with the same meaning as 'utterance'.

²⁰ Gumperz' definition of conversational involvement, a term originally suggested by Goffman (1961), is: 'Knowing what is going on', in terms of the social activity which the speaker and listener are engaged in, and being able to 'maintain, control and evaluate conversation' (1982a: 167, 1982 b: 16).

²¹ Gumperz defines intent in Gricean terms as 'the effect that a speaker intends to produce on a receiver by means of a message' (1982a: 156. See also Grice 1957, 1975).

²² Gumperz himself uses the terms 'prosodic and paralinguistic features': His 'prosodic features' are length, loudness, pitch, pause and the forms of intonation, while his 'paralinguistic features' correspond to what is referred to here as overall loudness, pitch, tempo, fluency and the voice quality *settings* (See discussions in Chapters 2 - 5).

²³ It should perhaps be pointed out that Gumperz' concept of (conversational) *involvement* is different from Tannen's concept of *involvement*, to be discussed in Section 1.2. below, and which may be glossed as *closeness*. This difference is discussed in some detail in Tannen 1989.

²⁴ See also 1982a: 162 ff. Gumperz' concept of the speech activity, adopted from Levinson 1978 and defined as 'the basic socially significant unit of interaction', corresponds to what is referred to as the *frame* of an utterance in anthropology, sociology and semantics (See Bateson

on a more local level, he must also signal the immediate communicative function of each utterance, referred to as the *discourse* or *communicative task*, examples of which are ‘explaining’, ‘narrating’ and ‘instructing’ (1982a: 208)²⁵. And thirdly, as a prerequisite for the signaling of both speech activities and discourse tasks, the speaker must also communicate the internal structure of individual sentences and their relationship to his message as a whole, a phenomenon generally referred to in linguistics as *discourse cohesion*. While discourse cohesion is also signaled by lexical items and syntactic structures, Gumperz is again particularly interested in the part that the suprasegmental features have to play in this process, as prosodic, or suprasegmental, *cohesion markers*.

1.1.1.1 The signaling of discourse cohesion - the suprasegmental cohesion markers

In the signaling of discourse cohesion, the first step towards clarifying a speaker’s communicative intent, the internal logical - or informational - structure of the speaker’s message is conveyed to the listener. This involves communicating the internal structure of individual sentences, distinguishing between items which are highlighted and others of lesser importance, and signaling their relationship to the other sentences within the speaker’s message²⁶. As Gumperz observes:

The first step in [the] inferential process is the participants’ perception of what is highlighted, what previous bits of information it is most likely related to, and how it is related. It is on the basis of such perceptions that participants [...] arrive at the interpretation of communicative intent and trace the line of argument (1982a:118). [...] For discourse to be cohesive, speakers must signal and listeners must interpret (1) what is the main part of a message and what is subsidiary or qualifying information, (2) what knowledge or attitudes are assumed to be shared, (3) what information is old and what is new, and (4) what is the speaker’s point of view and his/her relationship to or degree of involvement in what is being said. (1982b:28f.)

In his discussion of the suprasegmental cohesion markers Gumperz focuses on the role of length, loudness and pitch and pause, whose contribution - through the intonational processes of *tonality*, *tonicity* and *tune* - consists in the signaling of intonation-group boundaries, the marking of the *focus* and the creation of the intonational tune, or *contour* (1982a: 109f.)²⁷. While intonation-group boundaries divide the speaker’s message into short, easily absorbable units of information, corresponding to individual intonation-groups, the *focus* of such a unit constitutes its most important point. Finally, the intonation contour contributes to discourse cohesion by signaling the relationship between adjacent information-units, as for example whether the first leads onto the second or whether both units are to be seen as independent of each other.

1.1.1.2 Contextualization - the contextualization cues

Once the internal structure of the speaker’s message is clear, the listener can direct his attention to the *discourse task* and the *speech activity* the speaker is signaling. As both phenomena are intrinsically bound to each other, certain speech activities conditioning the occurrence of particular speech tasks and vice-versa, the knowledge of the speech activity is of utmost importance for the correct identification of the speech task(s). To name two examples, the knowledge that the speaker is teaching a French class (a speech activity) is important for the correct interpretation of a question like, ‘What did we do last week?’, to which an answer like,

1972, Hymes 1974, Goffman 1974 and Fillmore 1975, 1982). Tannen, who also uses the term *frame*, defines it as ‘a culturally determined, familiar activity’ (1986,1993). Also proposing the gloss ‘instructions for use’ (1986: 85), she notes that ‘in order to interpret utterances in accordance with the way in which they were intended, a hearer must know what ‘frame’ [the speaker] is operating in, that is, whether the activity being engaged in is joking, imitating, chatting, lecturing, or performing a play [...]’ (1993: 18).

²⁵ See also Gumperz 1982b: 11f., where the discourse/communicative task is glossed as a ‘goal-oriented recurring activity’. While on the surface the discourse task seems very similar to what is known as the ‘illocutionary force’ of utterances, in fact it is slightly more subtle in meaning (See Austin 1962, Searle 1969, 1975)

²⁶ For a detailed discussion of this phenomenon, also referred to as discourse coherence, see Halliday & Hasan 1976. Tannen 1984 also includes a discussion of discourse cohesion.

²⁷ The intonational phenomena of *tonality*, *tonicity* and *tune* (originally *tone*) which Gumperz adopts from Halliday (1967b, 1970), are discussed in detail in Chapter 2 below.

‘We had a fire practice’, would normally be inappropriate. Likewise, the knowledge that the enquiry, ‘Have you eaten?’, is a common Chinese greeting and thus not always meant literally as a question or an invitation, enables the competent speaker of Chinese to respond politely and correctly by always answering in the positive. It should however be pointed out that, as Gumperz emphasizes, knowledge of the speech activity does not actually determine utterance meaning, but merely serves to guide the listener towards the right interpretation²⁸. Thus, the knowledge that the speaker is teaching a class does not reveal the precise meaning of the above question, i.e. whether it was intended as the beginning of a relaxed group discussion or of an oral examination. Likewise, even if we know that our Chinese interlocutor *may* be engaged in the activity of greeting, we cannot *a priori* exclude the possibility that he might mean the question literally or wishes to invite us to dinner. In many interactions, especially in the context of intercultural communication, what exactly is meant, or intended, by the speaker is revealed slowly in the course of the conversation. We shall return to this point below.

In Gumperz’ interaction theory, the linguistic means by which speech activities and speech tasks are signaled are the *contextualization cues*, defined as ‘[features] of linguistic form that contribute to the signaling of contextual presuppositions’ (1982a:131). This definition covers a number of seemingly diverse speech phenomena, such as code-, dialect- and style-switching processes, idiomatic and formulaic expressions and the suprasegmental features. What they all have in common is their non-referential, indirect (and therefore ambiguous) meaning, which is highly dependent on the socio-communicative context of the conversational event in which they occur²⁹. Gumperz gives a number of examples which illustrate the implicitness and ambiguity of the contextualization cues. The first two involve formulaic speech in Black English and American English.

Git me a gig

Following an informal graduate seminar at a major university, a black student approached the instructor, who was about to leave the room accompanied by several other black and white students, and said: ‘Could I talk to you for a minute? I’m gonna apply for a fellowship and I was wondering if I could get a recommendation.’ The instructor replied: ‘Okay, come along to the office and tell me what you want me to do’. As the instructor and the rest of the group left the room, the black student said, turning his head ever so slightly to the other [black] students: ‘Ahma git me a gig!’ (Rough gloss: ‘I’m going to get myself some support’) (1982a: 30)

According to Gumperz, the black student’s use of the formulaic phrase *Git me a gig* was later identified by his fellow black students as a discourse strategy signaling the discourse task of ‘apology’ or ‘justification’ and meaning something like: ‘I’m still in control. I’m just playing the game as we blacks must do if we are to get along in a white-dominated world’ (1982a: 32).

Who’s the artist

A house painter arrived at the home of a middle class couple in California and was taken around the house to survey the job he was about to perform. When he entered a spacious living room area with numerous framed original paintings on the walls, he asked in a friendly way, ‘Who’s the *artist*?’ The wife, who was British, replied, ‘The painter’s not too well-known. He’s a modern London painter named ---.’ The house painter hesitated and then, looking puzzled, said, ‘I was wondering if someone in the family was an artist.’ (1982a: 144 - 145)

As Gumperz explains, this misunderstanding occurred because the British-born housewife was unaware that the house painter’s question was an American formulaic phrase. Whereas she interpreted it as a sign of the painter’s interest in art, it was in fact meant as a complimentary comment of the kind often uttered in America by a visitor who is seeing his host’s house for the first time (1982a: 145). While the woman interpreted the painter’s speech activity as ‘discussing modern art’, the activity he was in fact engaged in may be glossed as ‘a friendly

²⁸ In Gumperz’ words, the speech activity ‘channels [the listener’s] inferences by making relevant certain aspects of his background knowledge and underplaying others’ (1982a:131).

²⁹ As Gumperz notes, ‘Although such cues carry information, [their] meanings are conveyed as part of the interaction process. Unlike words, which can be discussed out of context, the meanings of contextualization cues are implicit [and] not usually talked about out of context’ (1982a: 131).

chat as part of getting to know and complimenting one's future employer'. This example shows how ambiguous the contextualization cues can be, even in conversations involving native speakers of what would normally be regarded as one and the same language. As Gumperz observes:

The mechanisms by which [contextual] information is signaled differ from lexicalized signs in one important respect: [...] they are inherently ambiguous, i.e. subject to multiple interpretations. In conversation such ambiguities are negotiated in the course of the interaction, through the manner in which second speakers respond to what they hear and through the reception that their countermoves receive. (1982a: 208)

1.1.2 The inferential process - the interpretation of speaker intent

The inferential process, during which the listener interprets the contextualizing information offered by the speaker, is defined in broad terms as 'the situated and context-bound interpretation by means of which participants in an exchange assess others' intentions and on which they base their response' (1982a: 153). Gumperz describes this process as follows:

We use our knowledge of grammar and lexicon, along with contextualization conventions and whatever background information we have about participants, to decide what discourse task is being performed and what activity is being signaled.[...] [Speech activity and discourse tasks] provide information about likely communicative goals and outcome. We then build on these predictions to identify the communicative intent that underlies particular utterances (1982a:18).

Due to the implicitness and ambiguity of the contextualization cues and the resulting complexity of the contextualization process, however, the interpretation of speaker intent through the identification of speech activities and discourse tasks is not always accomplished after one or two exchanges but may in fact extend over a quite large part of a given conversation, until the participants are quite certain that they are 'on the same wavelength' (1982a: 151). As Gumperz stresses, the success of this process, which then leads to the establishment of conversational cooperation, depends on the participation of all the interactants, who signal, through their responses and/or the backchanneling signals³⁰ they give, their view of what speech activity is being enacted and what conversational goals are being pursued. As he writes:

'[...] I would like to suggest that the signaling of speech activities is not a matter of unilateral action but rather of speaker-listener coordination involving rhythmic interchange of both verbal and nonverbal signals. [...] A successful interaction begins with each speaker talking in a certain mode, using certain contextualization cues. Participants, then, by the verbal style in which they respond and the listenership cues they produce, implicitly signal their agreement or disagreement; thus, they 'tune into' the others' way of speaking. Once this has been done, [...] both participants can reasonably assume that they have successfully negotiated a frame of interpretation, i.e. they have agreed on which activity is being enacted and how it is to be conducted' (1982a: 167)

The 'step-by-step' nature of the inferential process is illustrated by Gumperz' next example, in which the contextualization cues are overall pitch, loudness and tempo.

Joking on the plane to Miami

[The] incident was recorded when I was sitting in an aisle seat on an airplane bound for Miami, Florida. I noticed two middle-aged women walking towards the rear of the plane. Suddenly, I heard from behind, 'Tickets, please! Tickets, please!'. At first I was startled and [wondered] why someone would be asking for tickets so long after the start of the flight. Then one of the women smiled toward the other and said, 'I *told* you to leave him at home.' I looked up and saw a man passing the two women, saying: 'Step to the rear of the bus, please'. (1982a: 161)

As Gumperz recalls, his realization that the passengers were joking (a speech activity) developed only slowly, as the result of a gradual inferential process which built upon the cues he received little by little from their conversation. According to his account, the contextualization cues contributing most importantly here were the tempo and rhythm³¹ of the speakers' utterances as well as their overall pitch and loudness. They not only contextualized the male

³⁰ Backchannel signals, or listenership cues, are short responses, such as 'yes', 'oh' or 'mmh', which listeners often utter without intending to interrupt the current speaker. (Yngve 1970, Duncan 1973, 1974). While the use and meaning of these signals has been found to differ between different socio-cultural groups, including gender (Tannen 1993c, 1994, Coates 1998), most immediately and basically, back-channel signals convey that the listener is paying attention and is able to follow the speaker's line of reasoning.

³¹ By *rhythm* Gumperz means the way in which the tempo and timing of speaker and listener utterances may be said to harmonize, or cohere.

passenger's utterance as an imitation of an English bus conductor's typical call for tickets to be shown, but also indicated that the woman's following remarks were not meant literally (1982a:162). The following final example, involving the contextualizing functions of intonation contour, not only shows the ambiguity of the contextualization cues but also illustrates how their meaning, in this case the discourse task, may be interpreted quite differently by members of different speech communities.

I don't wanna read

In a taped elementary school classroom session, the teacher told a student to read. The student responded: 'I don't wanna read.' The teacher got annoyed and said: 'All right then, sit down.' (1982a: 147)

As Gumperz goes on to explain, white judges to whom this passage was presented for interpretation saw the child as being disobedient and impolite, signaling refusal. In contrast to this, speakers of Black English called attention to the rising intonation at the end of his utterance, which to them indicated that the student was asking for encouragement. In their view, the child was enacting the discourse task of a 'pleading request', meaning: 'Push me a little and I'll read. I can do it, but I need to know that you really want me to do it' (1982a: 147).

As a final point to be noted here, Gumperz emphasizes that any interpretation the listener arrives at, far from being definite, is initially a tentative and temporary hypothesis, liable to revision at any time:

The judgements made at any one time are contingent judgements. They are either confirmed or disproved by what happens subsequently. If they are confirmed, our expectations are reinforced; if they are disconfirmed we try to recode what we have heard and change our expectations of goals, outcomes and speakers' intent (1982b:18)

Therefore, the establishment of conversational involvement must always be followed by the maintenance of conversational involvement, the interactants checking from time to time that they still understand each other correctly. As Gumperz writes: 'Speakers continue in the same mode, assigning negotiated meanings to contextualization cues, until there is a perceptible break in rhythm, a shift of content or cues, or until a mismatch between content and cues suggests that something has gone wrong'(1982a:167). Thus, returning to the Chinese greeting, 'Have you eaten?', it is important that the listener, after answering that, thank-you, he has eaten, make quite sure that his interpretation of the speaker's question was correct. Over the next two or three exchanges he will therefore carefully listen for signals which might indicate that the speaker is not simply greeting him but in fact intends to invite him to a meal³².

While the temporary nature of any interpretation of speaker intent is largely due to the implicitness and ambiguity of the contextualization cues, it is also linked to the fact that people's conceptions of discourse tasks and speech activities often vary quite considerably³³. This is very much the case between members of different socio-cultural and ethnic groups, a point to which we come now.

1.1.3 The role of linguistic and socio-cultural background knowledge

The above discussed examples illustrate clearly that in order to interpret correctly what he has heard, a listener needs a certain amount of background knowledge which should be either identical or at least compatible with that of the speaker. In Gumperz' view both linguistic and socio-cultural background knowledge is needed for the correct interpretation of the contextualization cues. Linguistic background knowledge, relating to the use and interpretation of these cues, is defined as 'the knowledge of linguistic and related communicative conventions which speakers need to establish and sustain conversational cooperation, [comprising] both

³²As a few examples of such cues, the speaker might repeat the original question, ask where or what exactly the listener has eaten, or remark that he himself has not had anything yet, this possibly followed by the invitation to 'join him'.

³³It should perhaps be noted that discourse tasks, speech activities and even speaker intent may change during one and the same interaction, yet another fact which renders the correct interpretation of speaker intent very difficult (Gumperz 1982a: 131).

grammar and contextualization (1982a:209). Socio-cultural back-ground knowledge may be glossed as the collective body of socio-cultural knowledge, or knowledge of the world, which members of a given speech community acquire as a result of their interactive and conversational experience. According to Gumperz, this knowledge is stored in the mind in the form of 'culturally specific speech activities' and activated in the interpretation process, when the listener determines which speech activity is being signaled by the contextualization cues and judges how what he hears may be 'integrated into what he knows about his culture and the immediate [communicative] situation' (1982a: 207). Thus, returning to the Miami-bound plane, Gumperz' socio-cultural knowledge also played an important role in his interpretation of the exchange, relating not only to what an English bus conductor normally says and does as part of his everyday work routine, but also to the fact that tourists on their way to Miami often engage in such joking (1982a: 162).

After this outline of Gumperz' communication theory, we now briefly turn to two further aspects of his work of direct interest to the present study: his approach to intercultural communication and the link Gumperz sees between language and social identity.

1.1.4 Intercultural communication

In Gumperz' interaction theory intercultural communication is defined as conversational interaction between people speaking the same primary language but stemming from different cultural and/or social origins. Gumperz' approach to the study of intercultural communication emphasizes the strong link between a person's socio-cultural or ethnic provenance and the communicative conventions he adheres to, especially those regulating the suprasegmental features of speech. As he stresses, even if an immigrant or a member of an ethnic minority learns to speak the language of his adopted country perfectly fluently, the influence of the socio-cultural and communicative conventions of his native culture is normally so strong and so long-lasting that they continue to shape the manner in which he communicates with other people. Depending on the degree to which an immigrant family continues to identify with its native culture, the communicative habits of the original culture may even subsist through several generations. In Gumperz' view, many if not most of the misunderstandings of speaker meaning and intent which occur in intercultural communication are due to the incompatibility of the interactants' socio-cultural backgrounds and communicative habits, the latter especially with respect to the suprasegmental contextualization cues. The following example from Tannen (1986) shows how differences in communicative habits, or *habitus*³⁴, concerning the use of the contextualization cues, in this case overall pitch, can lead to not only misunderstandings in intercultural communication but also in misjudgements of speaker personality.

Extreme pitch-shifts

A Greek man married to an American woman accused her of speaking in an irritating monotone, especially when their tempers were strained. She felt terrible about this newly discovered failing and wondered why no one had ever mentioned it before. It never occurred to either of them that he found the tune of her talk monotonous because he was listening for the extreme shifts in pitch typical of Greek speakers, especially Greek women. [...] Her American habit of muting her expression of emotion when she was upset seemed unnatural to him (Tannen 1986: 51).

This example shows that, while the implicitness and ambiguity of the contextualization cues is also to blame, in intercultural communication things are rendered even more complex by the fact that, to a much stronger degree than in interpersonal communication, the use of these cues as well as their interpretation, in terms of speech activities, often varies strongly from one speaker to the other³⁵. As a third point, Tannen's example illustrates another unfortunate characteristic of the suprasegmental contextualization cues which has been omitted so far:

³⁴According to Thompson 1991, the *habitus* is 'a set of dispositions which incline agents to act and react in a certain way'. These dispositions generate practices, perceptions and attitudes which are regular without, however, being consciously co-ordinated or governed by any rule. (1991: 12, quoted in Couper-Kuhlen 1997: 168, where the concept is applied to communicative dispositions)

³⁵As Gumperz points out, this is all the more so as the interactants speak the same primary language and are therefore under the impression that they understand each other very well (1982a: 145).

under normal circumstances both their use and their interpretation are entirely subconscious. Thus, for example, the normal speaker of English does not consciously decide to raise the amplitude, speed and pitch on a word, phrase or utterance he intends to emphasize. Likewise, on the part of the listener, both the correct interpretation of these cues - in terms of emphasis - as well as their erroneous interpretation - in terms of irritation, for example - are equally subconscious. In fact, as Tannen emphasizes, these signaling processes are so deeply hidden that even if one of the interactants possesses the insightfulness to realize that something may have gone wrong in the communication process, it may still prove impossible for him to explain to the other what kind of mismatch occurred, and where or why. In addition to this, as the above example also illustrates, in intercultural communication misunderstandings of speaker intent involving the suprasegmental features all too often lead to misinterpretations not of referential meaning, but of his interpersonal attitude and even his personality. As Gumperz writes:

[The] speaker is said to be unfriendly, impertinent, rude, uncooperative, or fail to understand. Interactants do not ordinarily notice that the listener may have failed to perceive a shift in rhythm or a change in pronunciation. Miscommunication of this type, in other words, is regarded as a social faux-pas and leads to misjudgements of the speaker's intents (1982a:132). Even an odd word, or turn of phrase or a misunderstood tone of voice can seriously affect trust among participants (1982b: 8).

It goes without saying that the suspicion and distrust which easily develop out of such misunderstandings of speaker attitude are easily transferred to other members the speaker's socio-cultural or ethnic group. Thus, seemingly harmless misinterpretations of speaker meaning can give rise to prejudices and the cross-cultural stereotyping of social and ethnic groups and minorities, a point which is stressed again and again in interactional sociolinguistics.

1.1.5 Language and social identity

As a further consequence of the strong bond between the social, cultural and ethnic origin of an individual and his communicative habitus³⁶, speakers may be said to enact or indeed, as Gumperz sees it, *produce*, their social identity by the manner in which they use language in conversational interaction. As this includes the use of the contextualization cues, it follows that not only the use of phonological and stylistic variables but also that of the suprasegmental contextualization cues can serve to signal a speaker's affiliation to a certain socio-cultural group. Though at first sight this may not seem such a great misfortune, both Gumperz and Tannen have demonstrated that far from simply relating to speech community membership, misinterpretations of 'social identity' based on speakers' communicative habits all too often involve - most often negative - judgements of their personality. We will return to this bond between language and social identity below.

³⁶The socio-culturally patterned communicative *habitus* of a speaker corresponds to what Tannen calls 'conversational style'. See below.

1.2 Deborah Tannen: Interaction strategies and the ambiguity of the metamessage

Deborah Tannen's early work within the paradigm of interactional sociolinguistics focused on the problems encountered in interactions between Americans of different ethnic origins. The discoveries made in the course of this research, relating to the above mentioned bond between language and social identity, led to her profound interest in the signaling of speaker attitude. While agreeing with Gumperz that the establishment of conversational involvement through contextualization is the most important step towards pleasant and successful communication, Tannen places less emphasis on the communication of discourse tasks and speech activities, relating to *what* the speaker is doing, and focuses more on the communication of information signaling *why*, that is for what interpersonal, attitudinal, reasons, the speaker is pursuing a given activity. In her view the signaling of this attitudinal information, which she refers to as the *metamessage*³⁷ of speech, is perhaps an even more important part of the contextualization process than the signaling of speech activities and discourse tasks³⁸. As she observes :

Very little of what is said is important for the information expressed in the words. But that doesn't mean that talk isn't important. It's crucially important, as a way of showing that we are involved with each other, and how we feel about being involved. Our talk [says] something about our relationships. Information conveyed by the meaning of words is the message. What is communicated about relationships - attitudes toward each other, the occasion, and what we are saying - is the meta-message (1986:29).

The metamessage of speech is communicated through two different, yet intrinsically linked, processes: On the higher, interactional level, it is conveyed by the speaker's interactional behavior, relating to his manner of approaching and interacting with others and described in terms of different *politeness*, or *interaction strategies*. On the lower, linguistic level, the meta-message is also signaled by the speaker's habitual use of language in the enactment of the interaction strategies. In Tannen's approach to interaction, a speaker's socio-culturally patterned habitual use of language in conversation, is referred to as his or her *conversational style*. The link between these two levels resides in the fact that the choice of an interaction strategy determines the use and therefore also the meaning of the linguistic phenomena he employs in conversation. Similar to Gumperz, Tannen is particularly interested in the contribution of the suprasegmental features to the signaling of speaker attitude. In her view, the non-referential, indirect and therefore ambiguous nature of features such as overall loudness, pitch, tempo and fluency³⁹ are to blame for the fact that metamessages are often misunderstood, with dire consequences not only for communication but also for relationships.

Thus, Tannen expands Gumperz' communication theory by the following three points: First, the metamessage, as a new component of speaker meaning and intent, signaled as part of contextualization; Second, the politeness, or interaction strategies⁴⁰, which signal different speaker attitudes, or metamessages; And thirdly, conversational style, a speaker's manner of using language to enact a given interaction strategy and signal its metamessage. In her approach to interpersonal and intercultural communication, Tannen focuses on the question of how the underlying attitudinal disposition of a speaker may be said to influence the way he interacts with other people. In her view, not only does speaker attitude motivate the choice of interaction strategies, which in turn determine the use of the suprasegmental features, but the underlying attitude of a speaker is in fact communicated to the listener by the very manner in which he negotiates interpersonal communication by means of the interaction strategy he has chosen⁴¹. In the following, we first take a closer look at the metamessage with its two com-

³⁷Tannen adopts the concept of the metamessage from Bateson (1972), who is quoted as saying that 'no message can be interpreted except by reference to a superordinate message about how the communication is intended (Tannen 1984: 23).

³⁸Tannen calls these 'large' and 'small frames' (1986: 83f.).

³⁹Tannen refers to these as linguistic signals or cues. Her research focuses mainly on overall pitch, loudness, tempo and the use of pauses, but also includes intonational pitch-range, the pitch covered by nuclear and prenuclear pitch-patterns.

⁴⁰Tannen uses both terms. For reasons to be clarified below, the term interaction strategy will be preferred here.

⁴¹In a way, therefore, an interaction strategy is conceivable as a particular kind of contextualization strategy, as it contextualizes the speaker's message in terms of the underlying interpersonal attitude of the speaker.

ponents before coming to a more detailed discussion of Tannen's three interaction strategies, to which a fourth will be added below. The ensuing discussion of the four interaction strategies will focus on their different uses in the Western and the Eastern cultural spheres. After a brief discussion of Tannen's concept of conversational style, the chapter concludes with the ambiguity of the metamessage and its sociological and linguistic causes.

1.2.1 The metamessage of speech

In Tannen's interactional model, the interpersonal information which is communicated by the metamessage relates to two different aspects of human interaction: First, the *affinity* which the speaker feels for his interaction partner, described by Tannen in terms of his wish for *involvement* or *independence* and, second, *footing*, the power position in which the speaker sees himself with respect to the listener.

1.2.1.1 Involvement and independence

According to Tannen, much of how people interact with each other is determined by their fundamental need for involvement and independence. As she writes:

We need to get close to each other to have a sense of community, to feel we're not alone in the world. But we need to keep our distance from each other to preserve our independence, so others don't impose on or engulf us. This duality reflects the human condition. We are [both] individual and social creatures. We need other people to survive, but we want to survive as individuals (1986: 31).

Thus, people feel both the need to get close to, or involved with each other, as well as the need to keep their distance and preserve their independence. Unfortunately, our needs for involvement and independence are problematic for two reasons: First, they are conflicting needs which counteract one another, and second, far from simply alternating with each other, they are in fact simultaneous. Hence, the satisfaction of one need almost automatically and immediately leads to the feeling that the other need is being neglected and vice-versa, resulting in a never ending battle between one need and the other. Tannen's illustration of this state of affairs is the philosopher Schopenhauer's parable of porcupines hibernating together:

The porcupines huddle together for warmth, but their sharp quills prick each other, so they pull away. But then they get cold. They have to keep adjusting their closeness and distance to keep from freezing and from getting pricked by their fellow porcupines - the source of both comfort and pain (1986: 31).

Thus, one minute we signal involvement, as for example by agreeing to do someone a favour, by accepting their advice or their offer for conversation and by wearing a friendly expression and speaking in a warm, relaxed and focused way, while the next minute, feeling a sudden fear of being encroached or impeded on and losing our independence, we turn down an invitation to dinner, cut short the conversation, and signal by the preoccupied, or even impatient look on our face and a cooler, more distant manner of speaking, that we would rather be somewhere else right now. Again, no sooner are we on our own that we find ourselves worrying that we may have hurt the other person's feelings by sending them away. Thus, human communication is like a balancing act, since, in Tannen's words, 'no matter what relative value we place on involvement and independence and how we express these values, [we] are always balancing the conflicting needs for both' (1986:33).

Another point which complicates matters even more is the fact that human communication involves at least two people, both of whom have the needs for involvement and/or independence. Therefore, anything a person says or does to show he is honouring the interaction partner's need for involvement, or *closeness*, implicitly constitutes a threat not only to his own independence but also to that of his interaction partner. Likewise, any hint of independence, or *distance*, on his part is as much a rebuff to his partner's need for involvement as it is to his own. Thus, the interactional behavior of people who are accustomed to greeting others with what is intended as a warm hug and an affectionate embrace is often felt to be over-

bearing by people used to a more distant manner of approaching others. On the other hand, the behavior of people who for some reason (sometimes only temporarily) wish to keep to themselves and limit conversations to a polite minimum is easily misinterpreted by others in terms of cool distance or arrogance. Thus, intended warmth and closeness is easily misunderstood as an imposition and the tendency to dominate others, while polite distance is often misinterpreted in terms of superiority and arrogance⁴².

1.2.1.2 Footing - Power and solidarity⁴³

Apart from the speaker's wish for involvement/closeness and independence/distance, the metamessage also signals the *footing*, or 'power position', in terms of *equal* or *unequal*, which he feels he occupies with respect to his interaction partner⁴⁴. The relative *footing* of interactants often reflects their relative social standing. Thus, while superiors and subordinates are usually on unequal *footing*, subordinates are bound to each other by virtue of their *equality*. Tannen links the concept of *footing*, the relative power position of interaction partners, to the sociolinguistic principle of *power and solidarity*: If the relationship between interactants is unequal, this implies that in some way one of them has *power* over the other. On the other hand, if both partners are on equal standing, this evokes a feeling of equalness, i.e. *solidarity*. Thus, superiors and subordinates are separated from each other by power, while subordinates are bonded by *solidarity*. As simple as this may seem, the signaling of *footing* in human interaction is, however, beset with problems and ambiguities. Tannen's next two examples illustrate this:

Millie

Jack visits his grandmother in a nursing home. She boasts that she is really 'in' with the nurses because they call her [by her first name] Millie. Jack isn't pleased; he thinks they aren't treating his grandmother with proper respect. He feels the nurses are establishing toward her a footing that reinforces their position of power. She takes their using her first name as an expression of solidarity (1986: 101).

As Tannen explains, while Jack's grandmother understood the nurses' use of her first name as an expression of *solidarity* and equalness, he himself took it to be an expression of *power*, alluding to the inequality between caretakers and patients (1986: 101). How difficult it often is to distinguish between the signaling of *power* and *solidarity* is also illustrated by the next example:

Thanks, Mom

Two women were walking together from one building to another in order to attend a meeting. They were joined by a man they both knew who had just exited a third building on his way to the same meeting. One of the women greeted the man and remarked, 'Where's your coat?'. The man responded, 'Thanks, Mom.' (1994: 24)

According to Tannen, while the young woman meant the question as a show of friendly concern, in the service of *solidarity*, the male student took her remark as a gambit in a power exchange, evoking a mother telling her son to go and put his coat on (1994: 24).

As a final important point to be noted here, the *footing*, or 'power position', of interactants with respect to each other, in terms of not just unequal but *superior* or *inferior*, may not only be determined by their relative social status but can also be a reflection of their relative physical or mental capacities. Since all human beings have different interests and abilities, groups of people engaged in a common speech activity, such as discussing politics, reading a road-map or exchanging cake recipes, always contain one or two participants who are more

⁴²Lakoff (1975) and Scollon (1982) call this unfortunate state of affairs the *double bind* of communication (Tannen 1986: 34).

⁴³The sociolinguistic concept of *power and solidarity* was developed by Brown & Gilman (1960) who illustrate it with the help of second person pronoun use in French, German, Italian and Spanish. (See Fasold 1990) See Tannen 1994: 22ff. for a detailed discussion of *power and solidarity*.

⁴⁴The concept of footing was originally suggested by Goffman (1974), who also used the term *alignment*. It corresponds to what was referred to above as the *position*, or *standing*, of the speaker with respect to the listener. In the semantic differential model of Osgood et al. 1957, it corresponds to *potency*, the power relationship between the speaker and listener.

proficient than the others. In such cases, those who are found to be more competent are generally treated as holding a superior *footing*. Finally, in interactive situations of this kind, especially those involving participants who are unfamiliar with each other, the footing of individual participants is negotiated in a step-by-step process and established in the course of the interaction.

1.2.2 Interaction strategies

The interaction strategy was defined above in broad terms as a manner of approaching and interacting with others in social and conversational interaction. As such, it is related to the notion of politeness, which Tannen defines in the broader sense of ‘taking into account the effect of what we say on other people’ (1986:35). In sociolinguistics there exist a number of politeness models suggesting different politeness, or interaction strategies, serving to regulate human interaction⁴⁵. The model which Tannen adopts into her interaction theory is Robin Lakoff’s *Three Rules of Politeness* (1973)⁴⁶ which build upon the more basic model of Brown & Levinson (1978/1987). Given the continued importance of the concepts presented by these two scholars, especially for interpersonal interaction in and with the Eastern cultural sphere, the present discussion begins with Brown & Levinson 1978/1987 and their model of positive and negative politeness⁴⁷.

1.2.2.1 Brown & Levinson 1978/1987 - Positive and negative politeness

In Brown and Levinson’s model, politeness is understood in the sense of observing and attending to people’s needs for social dignity, or ‘face’⁴⁸. In analogy to the anthropologist Durkheim’s (1915) distinction between positive and negative religious rites, Brown and Levinson distinguish between two kinds of face need which correspond to the above discussed needs for involvement and independence. Positive face needs are the wish of every member [of a given community] that his qualities, wishes and needs be appreciated and shared by other members, while negative face needs are the desire of every individual that his actions remain unimpeded by others⁴⁹. Coming to the interactional means of being polite, Brown and Levinson differentiate between two kinds of politeness strategy: Strategies which attend to a person’s positive face needs are referred to as *positive politeness strategies*, those which serve to fulfill a person’s negative face needs are *negative politeness strategies*⁵⁰. Positive politeness strategies convey a feeling of involvement with and closeness to the interaction partner, while negative politeness strategies communicate the speaker’s respect for the partner’s need for distance and independence. Among the examples Brown and Levinson give for positive politeness strategies is the action of ‘noticing and attending to the interaction partner’s interests, wants, needs and/or goods’ (1987: 103 f.). As examples for negative politeness they cite the use of honorifics (1987: 181 f.) and the communication of regret and reluctance to impinge on the interaction partner, as for example with the apology ‘I hate to intrude, but...’ (1987: 187f.). In the following, Brown and Levinson’s positive and negative politeness strategies are referred to as *approaching* and *distancing strategies*.

⁴⁵ In the sociolinguistic literature the most widely used term is *politeness strategy* (Cf. Brown & Levinson 1978/1987, Goody 1978, Scollon 1995, Holmes 1995). Lakoff and Tannen, however, use the terms *politeness strategy* and *interaction strategy* alternatively. The latter term is preferred here, as the manner of approaching others in interaction often aims to fulfill the speaker’s wishes rather than those of the interaction partner, this giving rise to speech behavior which is far from considerate and polite.

⁴⁶ In Lakoff 1975 these three strategies are referred to as *Rules of Politeness*. In Lakoff 1979, where they are seen as part of a continuum, this term is changed to *Rules of Rapport*.

⁴⁷ Scollon’s (1995) politeness model will not receive separate treatment here. Though slightly different terms are used, the model presented by Scollon is basically the same as Lakoff’s.

⁴⁸ First proposed by Goffman (1967:5 ff.), the sociolinguistic concept of *face* is describable in terms of a person’s self-dignity or social image. In Scollon 1995 *face* is defined as: ‘the negotiated public image, mutually granted each other by participants in a communicative event’ (1995: 35). *Face* is discussed in Brown & Levinson 1987: 61 ff.

⁴⁹ Positive and negative face needs are described in Brown & Levinson 1987: 62 ff.

⁵⁰ There are many different ways of attending to positive and negative face needs. Brown & Levinson (1987) list 14 different strategies of positive politeness (1987: 101 ff.) and 9 negative politeness strategies (1987: 129 ff.).

1.2.2.2 Lakoff's Three Rules of Politeness (1973)

Robin Lakoff's *Three Rules of Politeness* consist of three strategies, *camaraderie*, *distance* and *deference*, the latter two representing two different types of distancing (negative politeness) strategies. Tannen (1984: 11) glosses Lakoff's three strategies as follows:

1. Camaraderie: *Be friendly*
2. Distance: *Don't impose*
3. Deference: *Give options*

According to Tannen, *camaraderie* corresponds to the cordial, direct and open behavior of people who treat their interaction partners as equals. Her illustration of this kind of interactional behavior evokes a person who, feeling thirsty, 'walks straight into our kitchen, throws open the refrigerator door and calls out: "Got any juice?"' (1984: 11-12). As Tannen points out, people displaying this kind of speech behavior usually expect others to follow the same rule and may feel surprised or hurt if this is not the case (1986: 36). In contrast to *camaraderie*, the strategies of *distance* and *deference* represent two different ways of not getting too close to others, of keeping a polite distance. The first, *distance*, involves the avoidance of imposing upon others. Thus, having been offered a drink, a person enacting this strategy will first decline, not wanting to cause inconvenience, but, if asked again may accept once he is convinced that this is not the case and the other person is quite sincere about his offer. *Deference*, on the other hand, relates to interactional behavior in which the speaker sees himself as inferior to the listener and therefore transfers the power of action or decision to him. In this case, a person asked what he wishes to drink is likely to say "Anything will be fine" or "I'll take what you're having", leaving the decision of the beverage up to the other (1986: 36). It is important to note that *distance*, like *camaraderie*, treats the interaction partner as equal, i.e. on *equal footing*, while in *deference*, the interactants are on *unequal footing*: Here the speaker treats his interaction partner as superior to himself.

1.2.2.3 A fourth interaction strategy: Support

The three *interaction strategies* discussed as part of Lakoff and Tannen's interaction theory are the approaching strategy of *camaraderie* and the two distancing strategies of *distance* and *deference*. It was seen above that in *camaraderie* and *distance* everybody is treated as equal, while *deference* involves an unequal footing between interaction partners. An analysis of these three strategies based on their signaling of closeness/distance (*affinity*) and equality/inequality (*footing*) is shown in Fig. 1.

<i>Footing</i>	Approaching/positive strategies (closeness)	Distancing/negative strategies (distance)
<i>Equality</i>	Camaraderie	Distance
<i>Unequalness</i>	-	Deference

Fig. 1 The Interaction Strategies of *camaraderie*, *distance* and *deference*

Fig. 1 shows that while this framework differentiates between equal and unequal *footing* with respect to the distancing strategies, this kind of distinction is not provided for with respect to approaching strategies, as it does not contain an approaching strategy in which the interaction partners are on unequal footing. Such a strategy, to be called *support*, whose effect was illustrated in the above discussed example, 'Thanks Mom', is proposed here. The interaction strategy of *support*, or *supportiveness*, is thus defined as a strategy of positive politeness, or approaching strategy which attends to the interaction partner's positive face needs, i.e. his

need for involvement, while at the same time establishing an unequal *footing* between them. He may do this by offering him a present he knows will please him, by offering assistance in the pursuit of some goal, or by doing him some favour⁵¹. In so doing, the person attending to the other's (positive) face needs raises himself to a higher, more powerful, position than the receiver, who is seen in need of this favour or help, placed in the provider's debt and therefore in an inferior position. Thus, while the strategy of *support* originally intends to express the speaker's affinity and *solidarity* with the listener, in fact it easily creates a footing of inequality between interactants. As was demonstrated in the above example 'Thanks Mom', this raising of the speaker's power position with respect to that of the listener involves a certain danger: Quite understandably perhaps, rather than feeling gratitude for the speaker's assistance, the interaction partner often takes offence at being judged in need of help and treated as inferior. The problems that may arise from the application of *support* are illustrated once more in the following example.

The danger of support

During a heated debate in class, a female student, resenting the way the professor had dismissed as invalid the arguments of a male student whom she happened to fancy, offered an argument in support of what he had said. Having succeeded in convincing the professor of the correctness of the young man's position, she later left the lecturing hall with the happy feeling of having come to the support of one she felt an affinity to. She was therefore deeply shocked when he approached her afterwards to say that he was quite capable of taking care of himself and had not needed her help⁵².

Thus, while the female student had wanted to come to the young man's assistance and signal camaraderie, he obviously misinterpreted her demonstration of *support* as a show of *power* and *superiority*, all the more perhaps, as unlike him she had succeeded in convincing the professor of the correctness of his, i.e. the young man's, arguments.

To summarize, the interaction model proposed here contains four interaction strategies: the approaching strategies of *camaraderie* and *support* and the distancing strategies of *distance* and *deference*. While *camaraderie* and *distance* are based on the equal footing of the interactants, in *support* and *deference* the interaction partners are on an unequal footing. The four interaction strategies are shown once more in Fig. 2, with their respective metamessages:

<i>Camaraderie:</i>	<i>Closeness (involvement) and equal footing</i>
<i>Support:</i>	<i>Closeness (involvement) and unequal footing</i>
<i>Distance:</i>	<i>Distance (independence) and equal footing</i>
<i>Deference:</i>	<i>Distance (independence) and unequal footing</i>

Fig. 2 The four interaction strategies and their metamessages

As the next point in the present discussion, we come to the manner in which these interaction strategies are employed in different speech communities.

1.2.3 The interaction strategies in interpersonal and intercultural communication

As Tannen observes, people often vary considerably with respect to the interaction strategies they apply in a given interactional situation. While the choice of interaction strategy is also motivated by criteria relating to the interaction itself, such as the speaker's familiarity with and relationship to the interaction partner (i.e. familiar or unfamiliar, equal or unequal) and the formality of the interactional situation, to a large extent a person's choice of interaction

⁵¹The strategy of support proposed here roughly corresponds to Brown & Levinson's positive politeness strategy no. 10, noted in passing above, in which the speaker helps the listener obtain his wishes and needs (1987: 125). It should perhaps be added that the present set of four interaction strategies is not exhaustive, but represents the four patterns of interactional behavior regarded here to be of greatest import to communication between China and the West. Naturally, many more interaction strategies can be developed on the basis of Brown & Levinson's and Lakoff's models.

⁵²One cannot help but wonder whether the misunderstandings of speaker intent in 'Thanks Mom' and 'The Danger of Support' were not also related to the fact that in both cases the helping (superior) person was a woman, the assisted (inferior) person a man.

strategy is determined by the communicative norms and conventions of the speech community he belongs to. As Tannen's following example shows, differences in the habitual use of interaction strategies can lead to serious misinterpretations of speaker attitude.

The danger of distance

A famous speaker appears at a conference only long enough to deliver a lecture, after which he disappears. His colleagues murmur about how he thinks he's too important to waste time listening to other speakers. Actually, he escapes quickly from the conference because social interaction with those he does not know is painful for him. Not knowing how to approach others, he stands aside and averts his glance, [unintentionally] giving the impression of being unapproachable. (1986: 111)

Thus, the speaker's employment of the distance strategy, due to his shyness and inability to approach other people, was misinterpreted in terms of cool superiority and arrogance. The next example involves the different politeness conventions of England and America, two countries normally seen as belonging to the same cultural sphere.

An American in Britain

An American woman visiting England was repeatedly offended - even, on bad days, enraged - when Britishers (sic) ignored her in settings in which she thought they should pay attention [to her]. For example, she was sitting at a booth in a rail-road station cafeteria. A couple began to settle into the opposite seat of the same booth. They unloaded their luggage; they laid their coats on the seat, he asked what she would like to eat and went off to get it, she slid into the booth facing the American. And throughout all this, they showed no sign of having noticed that someone was already sitting in the booth. When the British woman lit up a cigarette, the American had a concrete object for her anger. She began ostentatiously looking around for another table to move to. [...] The smoker immediately crushed out her cigarette and apologized. This showed that she had noticed that someone else was sitting in the booth, and that she was not inclined to disturb her. But then she went back to pretending the American wasn't there, a ruse in which her husband collaborated when he returned with their food and they ate it. (1986: 42)

As Tannen explains, to the American woman, accustomed to the strategy of *camaraderie*, politeness required that strangers sharing the same booth at least exchange a few words, if only a conventional "Is anyone sitting here?". The British couple's strategy of *distance* therefore seemed extremely impolite to her. However, far from being inconsiderate, they were in fact being polite by not imposing: By not acknowledging her presence they were also freeing her from the obligation to acknowledge theirs (1986: 42).

If different communicative habits lead to misunderstandings among members of the same cultural spheres, it is to be expected that such misinterpretations are even more likely to occur in intercultural communication. Indeed, the many cultures of the world often have very different norms with respect to the interactive behavior regarded as considerate, or polite, in formal and informal situations, so that clashes between different systems of politeness⁵³ are highly common in intercultural communication. In most societies of the western world people treat each other as equal and politeness norms emphasize the positive face needs of the individual. Considerate and polite behavior therefore typically involves the strategies *camaraderie* and *support*, while *distance* and *deference*, if used at all, are reserved for formal situations. As was illustrated in the example above, the habitual use of *camaraderie* applies especially to American society, where people approach each other in a very open and cordial manner, readily exchanging physical contact and quickly developing a close emotional relationship. Even in situations where interactants have just met for the first time personal, even intimate, details of their lives are exchanged in a manner unthinkable in many other societies of the western world, such as Britain or Germany. On the other hand, assistance in America is immediately and generously offered, often on the part of complete strangers, and any resulting unequal footing of interaction partners is ignored. Indeed, as in most western societies, the very idea that the kind of support people offer each other as part of everyday life should disturb the fundamental equality between them is inconceivable to most Americans and incompatible with the American way of life.

⁵³What Tannen refers to as the 'politeness system' of a given culture or society, determined by its rules of human interaction, relates to the norms that determine which interaction strategy or strategies are used by its members in different speech situations.

In contrast to this typically western pattern of behavior, most eastern societies, such as China Japan and Korea, adhere to what may be described as totally opposite norms of politeness. Here, the unequalness of people and the negative face needs of the individual are emphasized and the most commonly encountered politeness strategies are *distance* and *deference* of varying degrees, depending on the degree of familiarity between interaction partners. The most well-known signal of *deference* is of course the bow, most widespread in Japan, where it is still part of everyday social etiquette. Though the formal, deep bow is also encountered in other eastern societies, in all but the most formal situations bowing there is now mostly slight and brief, and often replaced by a gentle nod of the head. In highly formal encounters between social unequals a number of additional physical movements and postures are used to signal a degree of *deference* and inferiority which to the western observer often seems exaggerated and unnatural. These include not only the lowering of one's eyes but also the shuffling of one's feet, the latter in particular signaling a high degree of help- or, indeed, haplessness, creating the impression that the subordinate is literally falling over his or her feet to please the superior. Coming now to the second typically Asian interaction strategy, *distance* is the general norm, even in everyday conversations. With the exception of people bound by long-term personal relationships, interactants generally adopt an attitude of strong *distance* and reticence. Conversation is limited to a polite minimum, physical contact is avoided at all times and details of the interactants' personal lives are never part of the conversation.

Given these two highly different politeness systems, it is not surprising that problems occur in intercultural encounters involving members of eastern and western societies. Often enough the differences in the 'polite' behavior of interactants are so strong as to seem unsurmountable, especially when, due to the incompatibility of their politeness systems, none of the interactants can find a 'reasonable explanation' for the other's behavior and each finds himself doubting his interaction partner's good intentions. A feeling for the difficulties which can arise in such situations is provided by Tannen's following example:

Ignored in Japan

An American man lived, as many Japanese do, in frightfully close quarters - a tiny room separated from neighbouring rooms by paper-thin walls. In this case, the walls were literally made of paper. In order to preserve privacy in this most unprivate situation, his Japanese neighbours acted as if no one else lived there. They never showed signs of having overheard conversations, and if, while walking down the hall, they caught a neighbour with the door open, they steadfastly glued their gaze ahead as if they were alone in a desert (1986: 43)

As Tannen notes, the American was first puzzled, thinking that he was being snubbed by his neighbours. After a while, however, he realized that far from being rude, they were being polite, by omitting to show involvement and not imposing on him (1986: 43)⁵⁴. The next example, recounting an incident observed by this writer in Taiwan, involves the strategy of *deference*.

Taking a taxi in Taipei

Two elderly Chinese gentlemen, having to all appearances been out to lunch together, had hailed a taxi and were now taking leave of each other on the side of the street, while the cab was waiting. After a lengthy and heated discussion as to who should take this cab and thus depart first, one of them finally started climbing into the taxi. Having already turned away, as I believed the incident was over, my attention was immediately attracted back to the scene by a new back-and-forth. Now, the gentleman standing on the pavement was trying to slip a bank-note to the driver of the taxi, clearly intending to pay the fare, an offer which the other man, virtually hanging in the cab door, was clearly not prepared to accept. In the end, the gentleman now seated in the taxi prevailed and the other put away his hundred-dollar note, bowing repeatedly as the taxi finally departed.

The two Chinese gentlemen were behaving according to the *deference* rule, each trying to place the other on a pedestal, first by inviting, indeed pressing him, to take the taxi first, and

⁵⁴ As another example of a *distance* strategy in Japan, traditional politeness norms hold that a person's entry to and departure from a location where other people are present involves a violation of their negative face and as such must be apologized for. To the visitor from the West entering a company lavatory it may seem strange that, upon entering and before leaving, company members say, 'Excuse-me', to those present.

declining to leave himself, and then by offering - or refusing - the hundred Taiwan dollars for the taxi fare⁵⁵.

Probably the biggest difference between western and eastern norms of interaction lies in the concept of *face* briefly introduced above, which receives great emphasis in most oriental societies⁵⁶. In fact in some countries, such as China, Japan and Korea, consideration of *face* determines virtually every detail of the speech behavior of interactants, not only in formal and semiformal encounters but also in interactions between people quite familiar with each other, such as spouses⁵⁷, friends and colleagues. To illustrate this with an example from everyday life, an employee in a Chinese company who has met with a problem, for example with a computer program, and wishes to ask a colleague for help, cannot ask him directly, but must put the matter forward indirectly, for reasons of *face*. As a direct request obliges the addressee to help, even if he realizes he is not up to the task, this can lead to problems, since his failure inevitably results in a loss of *face* which he will then blame on the petitioner. The indirect methods available to the employee with a problem include, for example, telling his colleague about it in a neutral, non-obliging manner and asking him whether he knows of anyone who could help. As this does not amount to a direct request, the other is under no obligation to assist him but is free to choose whether to help or recommend someone else if he realizes he was not up to the task. This indirect manner of handling the matter guarantees that the petitioner will receive assistance from someone competent enough to help, and also avoids the loss of *face* of people mistakenly asked for assistance⁵⁸. As unaccustomed at this behavior may seem to the uninformed western observer, it is a typical way in which such a situation might be dealt with not only in Chinese society but also across most of southeast Asia. The different manner in which situations of this kind are approached in China and the West may account for many of the interpersonal problems which occur on an everyday basis in joint ventures between Chinese and Western companies.

As the above example illustrates, the phenomenon of *face* also plays a role in the choice and enactment of interaction strategies in the East. While in Chinese and Japanese interactions between familiar interaction partners *camaraderie* and *support* are also employed, their enactment is careful, slightly reticent and to the westerner seemingly less warm, less genuine. With respect to *camaraderie*, this reticence reflects the interactants' consideration for each other's negative *face* needs, as too much closeness easily incurs an encroachment on a person's need for *distance*. On the other hand, the general reticence in matters of *support* is due to a more complex state of affairs. In these societies, especially in more traditional areas, the granting of assistance, in the form of a good deed or a favour, immediately raises the benefactor to a superior position, while the receiver falls to an inferior position and loses (positive) *face*. Thus, while the benefactor 'gains' (positive) 'face-points', these are, so to say, 'deducted' from the receiver's 'score', in effect his (positive) *face*. The receiver now 'owes' the giver these 'face-points' and is obliged to 'pay him back' with a similar 'good deed', failure to do so incurring a further loss of (positive) *face* on his part. Understandably, this state of affairs is often far from agreeable for the receiver: Not only may the benefactor ask for the settlement of the debt at any time, but there is also no generally agreed-upon system to evaluate the 'value' of debts and determine the type/amount of good deed(s) necessary to 'repay' them in full. As a result, a receiver may conceivably find himself asked to repay a debt a number of times, or by doing a favour whose 'value' in his eyes far outweighs that of the original favour he re-

⁵⁵ At the time, around the year 1986, 100 Taiwan dollars corresponded to approximately 2.5 US dollars.

⁵⁶ The Chinese concept of *face* is discussed at length in Goffman 1967, Scollon 1995 and Günthner 1993. See also Hu 1944, Yang 1945, Hong 1985 and Gu 1990. Chinese distinguish between two different kinds of *face*, *mian* and *lian*. While the two terms are very often used alternatively, in fact strictly speaking, only *lian* corresponds to the sociolinguistic concept of *face*, the public image of a person which is dependent on his correct and virtuous social and moral comportment. In contrast, *mian* involves a public image of a rather different kind, which rests upon a person's material possessions and social connections. The link between *face* and Chinese management and decision making is discussed in Smith & Wang 1996 and Yates & Lee 1996.

⁵⁷ As Holmes (1995: 22) notes, the Japanese wife is expected to treat her husband with *deference*. The same may be said of the traditional Chinese wife, though there does seem to be a stronger equality of the sexes in modern China - as compared to conditions in Japan.

⁵⁸ Indirectness as a conversational strategy is discussed in detail by Tannen 1986: 65 ff. See also Brown & Levinson 1978/1987.

ceived. Given the unpleasantness of this condition, it is not surprising that in China offers of assistance are sometimes avoided. In fact, assistance of strangers and even acquaintances may actually be refused - on the grounds that 'one does not them anything' and 'there is no debt to settle'⁵⁹. The fact that in Chinese society good deeds, favours and gifts involve the gain, loss and repayment of 'face-points' is shown in the following example from Taiwan:

Confused in Taiwan

A newly wed couple, he a young Taiwanese, she a student from Germany, were preparing to receive their guests at their wedding reception in a Taipei hotel. The first guests were already arriving when the husband suddenly realized that he had forgotten to bring something very important, the *li-bu* (literally, 'ceremonial cloth'), and immediately sent a few of the helpers out to buy one. The bride, who had never heard of this *li-bu* before, thought it was some kind of special garment she would be required to wear and expected to be called any minute to try it on. After the helpers had returned, however, no such thing happened and after while the young woman, puzzled, went to look for her husband and find out what was going on. She found him at the entrance to the restaurant, where a large table had been set up and the *li-bu* was being prepared for use. To her amazement it turned out that this was not a piece of cloth but a kind of ceremonial 'account-book', in which the pecuniary gifts of the guests were to be entered by the helpers, each guest being required to sign his name below the marked amount. When she later asked her husband why the guests were submitted to what seemed to her such humiliating treatment, he explained that this was a custom which ensured that when they were invited to a similar event by one of their guests in the future they would know what amount of money would be appropriate as a gift⁶⁰.

We now come to a brief discussion of the communication of the interaction strategies through the suprasegmental features, before turning to the last point of this chapter, the ambiguity of their metamessages.

1.2.4 The use of suprasegmental features in the signaling of interaction strategies

A speaker's choice of interaction strategy has a conditioning effect on his use of language in conversational interaction. In particular his use of the suprasegmental features, part of what Tannen calls *conversational style*, is motivated by his prior choice of one of the four interaction strategies, whose metamessage they are to signal. Tannen defines conversational style as follows:

Anything that is said must be said in some way, and that way is [conversational] style. In order to understand [the deeper meaning of] any words spoken, you need to know how the words are meant: is the speaker joking, scolding, being friendly or rude? The way that these intentions are communicated are the features of conversational style: tone of voice, pausing, speeding up and slowing down, getting louder and softer, and so on - all the elements that make up not only what you say but how you say it (1984:2).

Tannen, like Gumperz, takes a particular interest in the contribution of the suprasegmental features to these processes, focusing on overall loudness, pitch, tempo and pause, intonation and, to a more limited extent, voice quality⁶¹. In her efforts towards establishing a link between the choice of interaction strategy and conversational style, Tannen focuses on the two styles she discovered in the course of her early work mentioned at the beginning of this section⁶². These are the so-called *high-involvement style*, broadly describable as an intensive form of *camaraderie*, which she found to be typical for New Yorkers of east European, Jewish extraction, and the *high-considerateness style*, a distance strategy, which she found in her British and Californian speakers. According to Tannen, speakers of the *high-involvement style* are motivated by a strong wish for closeness and involvement with their interaction partners. Among the features most typical for this style are elevated overall pitch and loudness and a fast tempo, combined with the avoidance of pauses, both within and between speaker turns, serving to express the speakers' high degree of interest and involvement in the topic

⁵⁹This is reflected in the literal meaning of *dui-bu-qi*, often thought to be the Chinese equivalent of the English apology, 'I'm sorry'. The fact that it actually translates as: 'I have no means of settling this debt' may help to explain why this expression is not used in all the situations in which the English apology 'I'm sorry' is appropriate.

⁶⁰In the future repayment of this kind of debt either the same or, if possible, a larger amount of money is appropriate.

⁶¹With respect to intonation, Tannen focuses mainly on pitch-range (1984: 116 ff.). Her mention of voice quality is rare and limited to what appears to correspond to the phonatory *setting of breathiness* (1984: 86).

⁶²Tannen 1984: 144 ff. See also Tannen 1981 and 1994: 53 ff. The *high-involvement style* is also discussed in Fasold 1990: 71ff.

and their interaction partner(s). In contrast to this, speakers of the *high-considerateness* style are motivated by the wish to keep a polite, yet distinct, distance. Typical here are comparatively low pitch and amplitude, rather slower speech tempo and a higher incidence of pauses, all intended to express the emotional distance of the speaker with respect to topic matter and the interaction partner.

1.2.5 *The ambiguity of the metamessage*

The ambiguity of the interaction strategies, in terms of the intended metamessage, has already been noted on a number of occasions. Due to this ambiguity, the metamessage originally intended by the speaker is often misread by the listener in terms of a different, all too often negative, speaker attitude. Thus, in the above discussed example '*The danger of distance*', the polite, shy distance which was originally intended by the speaker was misunderstood in terms of cool superiority, signaling not politeness but cool distance and *arrogance*. Similarly, in the examples '*Thanks Mom*' and '*The danger of support*', the speaker's supportive speech behavior was interpreted in terms of one-upmanship and superiority, also associable with *arrogance*. Indeed, as Tannen demonstrates, through the above link between language and social identity, misinterpretations of speaker intent can also lead to misjudgements of speaker personality. According to the results of her research, speakers of the *high-involvement style* are often thought to be overly inquisitive, nosy and overbearing. This is illustrated in her following example.

Dinner with the FBI

[...] Lucy's sister had dinner with a young man she had recently met. He seemed rather reticent, but Carol did her best to keep the conversation going and show interest in him. At the end of the evening the young man said, 'It was nice having dinner with the FBI'. (1986:57)

The fact that misunderstandings of the metamessage cause listeners to misinterpret a speaker's attitude and personality, i.e. their 'social identity', is a point of great concern to Tannen. As she writes: 'Everyone is judging others and being judged by their ways of talking. If these ways reflect different habits and expectations [of communication], then people are continually misjudged and misunderstood. You try to be nice and are judged pushy; You try to be considerate and are judged cold; You try to make a good impression in a job interview or at a cocktail party and see that the other person is annoyed rather than charmed (1984: 4).

Some of the most serious misjudgements of speaker attitude and personality occur in misunderstandings of the *deference* strategy, where the speaker is suspected of falseness and hypocrisy, manipulateness and arrogance. The problems which can result from a misunderstood deference strategy are illustrated in the following example:

It doesn't matter

A group of German students were planning to hold a farewell party for a Taiwanese student who was about to return home after having completed her studies. When they had more or less finished their preparations for the event, they decided to ask the Taiwanese student, who up to this point knew nothing of their plans, which date would suit her best. As it turned out however, she was too busy with the preparations for her journey to have time for a party. Having tried in vain to persuade her to come, the German students finally gave up and sadly pointed out that she would thus be leaving without seeing all her friends one last time. At this, however, instead of apologizing for not coming to what had been meant to be her own farewell party, the Taiwanese student said, 'Oh, it doesn't matter'. This response confused the German students very much and left a number of them wondering whether her friendship had been genuine or whether she was in fact glad to see the last of them. It never occurred to them that what she had meant was: 'It doesn't matter if I'm not there. I'm not important. Go ahead and have a party. You don't need to cancel it if I can't come'.

Thus, while the Taiwanese student had intended to signal deference, by saying how unimportant her presence at the party was, this was totally misunderstood by the German students. Not accustomed to the employment of the deference strategy, they took her remark as a sign of cool distance and insincerity.

Very serious misinterpretations of speaker attitude may also occur when interactants do not share the same communicative habits with respect to humour, definable as speech which is

‘not meant literally’ (Tannen 1984: 130)⁶³. Thus, as Tannen shows, the self-mocking humour of the high-involvement style is often misunderstood in the opposite sense, as making fun of the listener (1986: 127f.). Possibly even more easily - and gravely - misunderstood is the form of teasing commonly referred to as *put-down* humour, which involves a playfully meant attack on the listener’s positive face by the speaker, who pretends to criticise or make fun of a certain aspect of the listener’s person - his or her looks or actions, physical or mental abilities. Depending on the type of humour involved, the listener may either just laugh or return the ‘attack’ by making a similar humorous remark about the speaker. While such ‘teasing matches’ are immensely pleasurable to people familiar with this type of humour, misunderstandings involving this kind of joking strategy, very popular in China, can have disastrous consequences for interpersonal relationships. This is illustrated in the following example.

The perils of humour

A female Taiwanese student who was studying at a German university on an exchange program was engaged in casual conversation with two students, one of whom was a young man from the Middle East, whom she had got to know quite well, or so she thought. Having heard him declare that in this beautiful weather he could not stay at home and study but was planning to spend the afternoon at the swimming-pool, she suddenly felt the urge to tease him and, without pausing to think, jokingly accused him of not going to the pool to swim but merely to relish the sight of so many scantily-clad young women, a banter which in her own culture would have prompted a well-phrased and equally teasing response. However, instead of returning the tease, the young man stared at her in contemptuous disbelief and left the room in a huff, mumbling something about the arrogance of judging others by oneself. Sadly, as a result of this misunderstanding, he staunchly avoided contact with her during the rest of her stay, and that was the end of what might have become a warm friendship.

Thus, whereas the young woman’s remark was meant as a ‘playful attack’ and an invitation to join the game and ‘shoot back’, the young man, not accustomed to this form of humour, took the remark seriously and believed he was being subjected to a very strong form of criticism. While her message was one of *camaraderie*, it was interpreted in terms of *arrogance*. This particular form of teasing, challenging, put-down humour is referred to here as ‘*tai-gang*’, the term in Chinese⁶⁴. The examples discussed above show that the attitude signaled by an interaction strategy is often ambiguous. We now come to the causes of this ambiguity.

1.2.6 The origins of the ambiguity of the metamessage

According to Tannen, the ambiguity of the metamessage has two main origins: On the one hand, the interaction strategies are highly ambiguous, due to the underlying sociolinguistic processes involved - *power and solidarity*. On the other hand, due to their non-referential, indirect and therefore highly ambiguous meanings, the suprasegmental features, which play an important part in the signaling of metamessages, are also to blame for misunderstandings of speaker meaning and attitude.

1.2.6.1 The sociolinguistic origins - power and solidarity

The sociolinguistic causes of the ambiguity of the interaction strategies are found in the ambiguity of *power and solidarity*, a process introduced above and linked to the concept of footing. As Tannen has demonstrated, all four interaction strategies are easily misinterpreted with respect to equalness and inequality of footing, i.e. *power* and *solidarity*. Thus, the strategy of *support*, meant to signal affinity and *solidarity*, is easily misunderstood as treating the listener as in need of help, and therefore inferior (a form of unequalness), while in *deference*, where the interaction partner is actually being raised to a higher position, he suspects the speaker of wanting to manipulate him, another form of treating him as inferior. As the above discussed

⁶³Irony and joking as part of conversational style are discussed by Tannen in 1986: 130 ff. For detailed discussions of irony, see Roy 1978 and Kaufer 1981. For discussions of humour in general, see Palmer 1994 and Ross 1998. A brief section on irony and sarcasm, to which reference will be made in Chapter 6, is included in Norrick 1993.

⁶⁴*Tai-gang*, translated as ‘to argue’ in the New Practical Chinese-English Dictionary, is discussed in more detail in Part Three below. It bears a strong similarity to what is described in the literature as verbal duelling: informal linguistic contests, in which people attack each other through forceful or ingenious use of language’ (Crystal 1987: 60). Tannen finds a similar joking strategy, which she refers to as ‘ritual complaining’, with speakers of the *high-involvement style* (1986: 58).

examples have shown, even the two interaction strategies meant to signal equalness and solidarity, i.e. *camaraderie* and *distance*, may be misunderstood in similar negative ways, especially in intercultural interactions. In fact, it often seems as though the ever suspicious listener, refusing to give the speaker the slightest benefit of the doubt, were constantly 'looking out' for the smallest indication of inequality/power in the speaker's speech behavior, holding it up almost triumphantly as 'proof' of the speaker's ill-will⁶⁵. As Tannen explains, the ambiguity of the interaction strategies with respect to *power* and *solidarity*, equality and unequalness, is due to the fact that, like the human need for involvement and independence, both dimensions are always present. As she writes: 'It's a paradox, like the drawing of a chalice and two faces. Both images exist in the picture simultaneously, and we can see both, but we can only see one at a time. In the same way, we can see only one side of the *power/solidarity* dimension at a time.' (1986: 102)

1.2.6.2 The linguistic origins - the ambiguity of the suprasegmental features

Gumperz and Tannen repeatedly call attention to the fact that, due to their non-referential, indirect and ambiguous meanings, the suprasegmental features can be used to communicate a large number of different messages, relating to speech activities, discourse tasks (in contextualization) and speaker attitude (metamessage). As Tannen cautions:

[The suprasegmental features] are used in devices that do the daily work of having conversations - work like showing you're listening, interested, establishing solidarity - or that you're not. Usually these devices work just fine, but because they're not explicit, they can be misinterpreted.' (1986: 54)

Thus, while the elevated loudness, tempo and pitch of the *high involvement style* serve to communicate the strong interest and involvement of the speaker with respect to topic matter and interaction partner, precisely the same signals can also be used to signal other, very different messages, such as negative surprise, frustration and even anger. Naturally, this can - and frequently does - lead to misunderstandings of speaker attitude and misjudgements of personality, examples of which were discussed above. Gumperz and Tannen also emphasize that, due to the implicitness and ambiguity of the suprasegmental features and the fact that their use in conversational interaction is regulated by socio-culturally patterned communicative conventions, listeners not possessing the appropriate background knowledge relating to these conventions are often unable to correctly interpret the meaning of these signals and therefore misinterpret their interpersonal messages.

⁶⁵ Therefore, it would seem that the strategy of *support* is almost predestined to be misunderstood, even within one culture, by virtue of the fact that assistance is usually offered by more competent persons to people less competent and in thus need. Another source of miscommunication, this time between China and the West, is the different manner in which *support* is offered in the two cultural spheres. While the Chinese system of *face*-'points' is inconceivable to most westerners, the western custom of 'free of charge' gifts and favours, i.e. without the incurrance of debts and the owing of favours is quite as difficult to understand for the average Chinese. While westerners are likely to distrust the Chinese custom as calculative and hypocritical, to most Chinese the western custom appears naïve - as well as inconsiderate, if the receiver is not given a chance to 'repay his debt', which in much of China amounts to a moral duty.

Summary of Part One

In interactional sociolinguistics the interaction partners must establish and maintain conversational involvement as a prerequisite to successful interaction. This is achieved through the process of contextualization, during which, with the help of the contextualization cues, the information needed to understand speaker meaning and intent is signaled. Given the implicitness and ambiguity of these signals, whose use is regulated by socio-culturally patterned conventions, the successful outcome of this process requires that the interactants possess a common, or at least compatible, form of linguistic and socio-cultural background knowledge. While Gumperz' primary concern is the communication of speech activities and discourse tasks, signaling in effect *what* the speaker is doing, Tannen's interest focuses on a different level of contextualization, involving *why*, i.e. for what attitudinally, i.e. interpersonally, motivated reason, the speaker is engaged in a given speech activity. In Tannen's interaction model, a speaker's interpersonal attitude is communicated in the form of the metamessage, which represents the attitudinal counterpart to Gumperz' more factually oriented contextualizing information. In Tannen's model the metamessage of speech is communicated, first, through the speaker's interactional behavior, in terms of the interaction strategy which his actions and utterances are intended to convey, and secondly, through the linguistic signals, in particular the suprasegmental features, which he uses in the enactment of this interaction strategy. Tannen sees the interpersonal-attitudinal information conveyed by the metamessage as consisting of two components: *Affinity*, described as the speaker's wish for *involvement closeness* or *independence/distance*, and *footing*, the speaker's *equal* or *unequal* position, or standing, with regards to the listener. In the expanded interaction model developed at the end of Chapter 1, the four interaction strategies *camaraderie*, *support*, *distance* and *deference* were defined in terms of the following metamessages:

- (1) *Camaraderie*: Wish for closeness/involvement and equal footing
- (2) *Support*: Wish for closeness/involvement and unequal (superior) footing
- (3) *Distance*: Wish for distance/independence and equal footing
- (4) *Deference*: Wish for distance/independence and unequal (inferior) footing

As Tannen demonstrates, the signaling and interpretation of the interaction strategies and their metamessages, far from being easy and straightforward, is a complicated process which often goes wrong, resulting in misunderstandings of speaker attitude and misjudgements of speaker personality. Speech behavior intended to signal *camaraderie* is often misunderstood as an overbearing and pushy attitude, while *deference* is often mistaken for falseness, hypocrisy and manipulateness. In a similar fashion, the *distance* strategy is easily misread in terms of cool or cold aloofness and arrogance, while *support* is mistaken for one-upmanship, superiority and arrogance. Fig. 3 lists the four interaction strategies with their intended metamessages and their *antagonists*, the attitudinal message in terms of which they are most commonly misunderstood:

<i>interaction strategy</i>	<i>intended metamessage:</i>	<i>misinterpretation (antagonist):</i>
<i>camaraderie</i>	closeness/equalness	pushiness
<i>distance</i>	distance/equalness	cold aloofness
<i>support</i>	closeness/inequality	arrogance
<i>deference</i>	distance/inequality	falseness, hypocrisy

Fig. 3 The interaction strategies with their metamessages and antagonists

The fact that misunderstandings of the metamessage give rise not only to misunderstandings of attitude but also, through the bond between language and social identity, lead the listener to misjudge the speaker's personality, is a point of utmost concern to Tannen. Given the unhappy consequences of such misunderstandings for interpersonal and intercultural communication, she takes a strong interest in their causes. According to the results of her research, these reside in the ambiguity of the underlying sociolinguistic processes (*power* and *solidarity*) and the ambiguity of the suprasegmental features.

Of great interest to both Gumperz and Tannen are the suprasegmental contextualization cues, the suprasegmental features, whose non-referential and ambiguous meanings they blame for the miscommunications which often occur in contextualization and the signaling of speaker attitude. Since the use of these linguistic signals is determined by socio-culturally patterned rules and conventions, both linguists stress that the correct interpretation of these signals is only possible for a listener who shares the speaker's linguistic and socio-cultural background knowledge and who therefore has knowledge of the speaker's communicative habits, in Tannen's model also referred to as his or her conversational style.

While Gumperz and Tannen are certainly right that many of the above discussed problems of interpersonal and intercultural communication may be traced to the indirectness and ambiguity of the suprasegmental features, their writings do not include detailed discussions of the origins and the precise nature of this ambiguity. These issues will be addressed in Part Two, which contains a phonetically oriented definition and description of the suprasegmental features, the communicative forms they give rise to and the communicative functions which these fulfill. The aim of the discussion will be to find out, on the basis of the research literature, what may be said about the role of the suprasegmental features, both in the signaling of communicative meaning in general as well as the communication of the metamessage of speech, as defined by Deborah Tannen. It is, for example, conceivable that the signaling and interpretation of the four metamessages relies on other suprasegmental features than those discussed by Tannen, and that the purported ambiguity of the metamessage is only due to her omission of these features. As one possible example, of such a feature, the role of voice quality in the signaling of communicative meaning will come under close scrutiny in Part Two.

Part Two

Dimensions of meaning in suprasegmental communication

Introduction to Part Two

The four chapters which make up Part Two are devoted to the *forms* and *functions* of suprasegmental communication in German - as a western, Indo-European, language - and Mandarin Chinese¹ - as a representative of the Sino-Tibetan tone languages. The (communicative) form may be defined as a linguistic phenomenon, such as a nuclear pitch-pattern, a certain degree of loudness or voice quality (setting), which is used to signal communicative meaning, via the creation of meaningful contrasts². The most important objectives pursued in the discussions to follow are: First, finding a phonetically explicit and at the same time functionally oriented definition of the forms of suprasegmental communication in German and Chinese; Second, determining the functions of these forms in human communication, i.e. the manner in which they are used to signal communicative, in particular attitudinal, meaning; And third, detecting the origins of the above mentioned functional ambiguity of the suprasegmental forms. Three groups of communicative forms are distinguished: (1) the forms of *intonation*, presented in chapters 2 and 3; (2) the *overall features*³ - overall pitch, loudness, tempo and fluency - discussed in chapter 4; and (3) the forms of *voice quality*, discussed in chapter 5. The basic building-blocks of these forms are the suprasegmental features *length*, *loudness*, *pitch*, *pause* and *voice quality*. Short and longer-term variations in the first four features give rise to the forms of *intonation* and the *overall forms*, while the different types, or *settings*, of voice quality correspond to variations in the speaker's voice quality. We begin with a brief presentation of the suprasegmental features with their physiological, acoustic and perceptual correlates and the manner in which they contribute to communicative information, by means of the communicative forms defined above.

Length

Defined in perceptual terms as the relative length - *short*, *mid* and *long* - of a speech sound in a given environment, the acoustic correlate of length is the *duration* of the corresponding speech signal, indicated in milliseconds (*ms.*). On the physiological plane length corresponds to the time required for the realization of the speech sound, dependent on articulatory speed, a function of the velocity of the articulators involved in its realization and the extent of the movements carried out by the articulatory organs.

Loudness

Defined perceptually as the impression that a given speech sound is louder, or softer than another in its environment, loudness corresponds to two acoustic measures, *amplitude* and *intensity*. The *amplitude* of a speech signal, indicated in Pascals (*Pa*), is defined as the maximal displacement of the air particles involved in its sound wave, reflected in the peak-to-peak values of its speech pressure waveform. The *intensity* of a speech signal, measured in decibels (*dB*), corresponds to the amount of energy present in the signal⁴. On the physiolo-

¹ In the following, Mandarin Chinese, the modern standard language of China, Taiwan and Hongkong, will be referred to simply as Chinese.

² This inherently functional concept of communicative meaning, based on the presence of 'meaningful contrasts', constitutes the fundamental point of departure of all modern linguistic research - from the *minimal pairs* of structural phonology to the componential analyses of semantic theory. Note, however, that while the concept of contrast is relatively unproblematic within core linguistics, where meaning is defined in terms of referential, factual information, the question of what exactly is to be regarded as meaning in suprasegmental communication is less straightforward, since much more than factual information is involved. According to the definition proposed by Lyons (1977), to be adopted here, a meaningful contrast between two 'minimally distinct intonational pitch-patterns' exists when native speakers associate different, either referential, expressive or social, meanings with different members of the pair, or when different contexts would be appropriate for the use of different members of the pair' (1977: 50 f.). This definition leads us to a further problem in suprasegmental communication which has yet to be resolved in a satisfactory manner, that of 'minimal distinctiveness', a term borrowed from the binary approach of segmental phonology. Naturally, the distinction between different forms of suprasegmental communication, such as varying degrees of loudness or different intonational pitch-patterns, is not *binary* but *gradual* and can therefore not be analysed in terms of the presence or absence of a given feature. (See further discussion in Couper-Kuhlen 1986: 110)

³ The terms 'overall' length, loudness and tempo of utterances, relating to speech phenomena based on longer-term variations of these features, were chosen to avoid confusion with the short-term variations in length, loudness and tempo, which give rise to the forms of intonation (See discussion below). The discussion of overall pitch will also include a brief account of (overall) pitch-range, which plays a minor role in Western languages but holds important functions in the Chinese language, in particular Chinese intonation (See Chapter 3).

⁴ Note, however, that while the perceived loudness of a (speech) signal is related to its intensity, these two parameters are not in perfect

gical plane, loudness is a function of the subglottal air-pressure created below the *larynx*⁵ by the respiratory system. It is also dependent on the degree of mouth opening, which modifies the amount of energy radiated during speech.

Pitch

On the perceptual plane, pitch is defined as the perceived pitch-height, e.g. *high*, *mid* or *low*, of a voiced speech sound as compared to others in its vicinity. Its articulatory correlate is the *frequency* of vocal fold vibration, modified by subglottal air-pressure and muscular adjustments in the larynx which regulate the length, mass and tension of the vocal folds. On the acoustic dimension, the pitch of a speech sound, corresponding to the excitation-rate of the air molecules involved, is defined as the *fundamental frequency* (*fo*) of its signal, calculated as the number of repetitions, or cycles, of its waveform per second, and given in Hertz (*Hz.*).

Pause

The pause is describable in broad terms as an interruption in a given flow of speech. Pauses may be either silent or 'filled': The silent, or *empty*, pause, is defined perceptually and acoustically as a period of silence⁶ and corresponds on the articulatory plane to the cessation of vocal activity. The filled pause contains an unspecified number of so-called *fillers*, single or repeated vocal phenomena, such as 'mmh', 'ah' or 'er'. A variant of the filled pause is the *lengthening*, defined as the prolongation of a syllable/segment beyond its normal, unmarked, duration, a phenomenon commonly encountered towards the end of utterances, as a result of final slow-down. While pauses and lengthenings are described acoustically in terms of their duration in milliseconds (*ms.*), perceptual assessments of pauses are generally made in terms of the relative length - *long*, *mid* or *short* - in comparison to the length of single syllables or segments in the immediate environment⁷. The perceptual, articulatory and acoustic correlates of *length*, *loudness*, *pitch* and *pause* are shown in Fig. 4⁸.

Voice quality

The *voice quality*, or *tone of voice*, of a speaker may be described in perceptual terms as a certain auditory coloring of his voice, is heard to be running through all the sounds of his utterances (Abercrombie 1967: 91)⁹. Acoustically, this auditory coloring may be described in terms of acoustic properties shared by these sounds. On the articulatory plane, the configurational approach of John Laver (1980) defines voice quality in terms of certain configurations of muscular adjustments, referred to as *articulatory settings*, and effected in the laryngeal or supralaryngeal part of the vocal apparatus. In Laver's model the *setting* is regarded as the basic descriptive unit and communicative form of voice quality communication.

correlation: Due to differences in the responsiveness of the human perceptory system, sensations of equal loudness for different frequencies are based on different intensities. To name another non-linear correlation in this context, the perceptory impression of loudness is also dependent on interactions between the intensity and duration of speech signals.

⁵ The larynx, sometimes also referred to as the *voice box*, is situated at the top end of the trachea and contains the vocal folds. (See Chapter 5)

⁶ Strictly speaking, a definition in terms of 'relative' silence would be more correct, as one has to take account of the level of the background noise which is usually present in recordings of natural speech - and therefore also in the spectrograms of this type of speech.

⁷ It should perhaps be added that the relative length of a pause as perceived by a listener is of course dependent on speech tempo: Depending on the speed with which an utterance is delivered, a pause of, say, 0.5 secs. may be judged perceptually as *long*, *mid* or *short*.

⁸ Fig. 4 is adapted from Couper-Kuhlen (1986: 7). As Couper-Kuhlen points out, the relationships between these three levels are not as straightforward as they might seem. Thus, perceptual judgements of the loudness, pitch and length of speech sounds are not linearly correlated with the intensity, frequency and duration of their corresponding acoustic signals. As a second point, the precise physiological correlates of pitch and frequency on the one hand, and loudness and intensity on the other, still await clarification. As a third point, the acoustic measures of intensity, fundamental frequency and duration are dependent on a number of other factors, such as the phonetic environment of speech sounds, vowel quality and the point of articulation of consonants. (Discussions of these influences are found in Couper-Kuhlen 1986: 7f and Cruttenden 1986: 2f.)

⁹ Segments often differ in the degree to which they are affected by, or susceptible to, the articulatory and perceptory properties of voice qualities, in particular of the *supralaryngeal settings*. The scale of the effect of a given supralaryngeal *setting* on a segment is proportional to the distance separating their articulatory locations. Thus, for example, velar consonants and open back vowels are highly susceptible to *palatalization*, brought about by moving the tip of the tongue towards the front part of the hard palate, whereas the effect of this *setting* on high, close front vowels and bilabial or dental consonants is minimal. (See discussion in Laver 1980: 20 f.)

Feature	Perceptual correlate	Physiological correlate	Acoustic correlate
<i>Length</i>	perceived duration of speech sound	Duration of articulatory process	Duration of speech signal
<i>Loudness</i>	perceived loudness of speech sound	Strength of pulmonary air-pressure	Amplitude/intensity of speech signal
<i>Pitch</i>	perceived pitch of speech sound	Vibration-rate of the vocal folds	Fundamental frequency of speech signal
<i>Pause</i>	perceived interruption in flow of speech	(1) cessation of vocal activity (2) articulation of <i>fillers</i>	(1) silence (2) acoustic signals of <i>fillers</i>

Fig. 4 The phonetic correlates of *length*, *loudness*, *pitch* and *pause*

Forms and functions in suprasegmental communication

Part of the ambiguity of suprasegmental communication to which attention was called in Chapter 1 is related to the manner in which the suprasegmental features, in particular length, loudness, pitch and pause, contribute to human communication. In contrast to voice quality communication, which is based on one communicative form, the *articulatory setting*, the manner in which length, loudness, pitch and pause contribute to communication is comparatively complex, since they do so through the action of not one but three different types of communicative forms, bound to different time-domains. Thus, local variations in length, loudness and pitch extending over a small number of syllables within one and the same utterance contribute - together with phonological pauses/lengthenings - to the creation and demarcation of the *forms of intonation*, while longer-term, overall, variations in length, loudness and pitch and non-phonological pauses/lengthenings give rise to the *overall features* of loudness, pitch, tempo and fluency. The third type of communicative form based on length, loudness, pitch and pause are the *long-term*, (quasi-) permanent, *loudness*, *pitch*, *tempo* and *fluency features* of a person's voice which can serve in the speaker identification and voice profiling¹⁰.

A second factor contributing to the complexity of suprasegmental communication is the fact that the majority of the communicative forms discussed here hold more than one communicative function. Thus, while the voice quality *setting* of *nasality* when limited to single segments¹¹ functions phonologically as a distinctive feature of nasal consonants and vowels, while the longer-term use of *nasality*, extending over whole utterances, serves to signal emotional and attitudinal information, such as scorn, contempt and sarcasm. Finally, *nasality* as a (quasi-) permanent feature of a person's voice can identify him or her as a speaker of a certain variety of English, such as RP¹². While the different communicative functions of voice quality *settings* are dependent on the functional domain (from single segments and syllables to whole utterances and even longer stretches of speech) and thus also the time-domain involved, the many functions of the communicative forms based on length, loudness, pitch and

¹⁰ As is the case with voice quality, these long-term features are for the most part physiologically determined, but can be altered to a certain degree with speech therapy (Graddol & Swann 1989). Otherwise, these features usually remain unchanged over long periods of time, though minor can occur as the result of illness, professional activity and periods of high emotional and psychological stress.

¹¹ Due to assimilatory processes often encountered in nasality, this *setting* can also affect the vowels and consonants in the immediate vicinity of nasal consonants and nasalized vowels, leading to pre- and postnasalization.

¹² The term Received Pronunciation (RP) relates to the regionally neutral, 'standard' accent of British English, originally that used at Court and in the public schools. The American equivalent is *General American*, the German standard is *Hochdeutsch* ('High German').

pause are quite independent of the time-domain. Thus, both the forms of *intonation* and the *overall features* can serve to signal *linguistic*, or *discoursal*, information, showing how consecutive utterances and speaker contributions belong together; As a second function, they can also communicate *paralinguistic*, i.e. emotional and attitudinal information; And as a third function, they can also serve to communicate *extralinguistic*, indexical, information about the speaker, such as his affiliation with certain socio-regional variants, his membership to particular professional groups and certain traits of his personality. As already mentioned above in the General Introduction, this functional diversity of the suprasegmental phenomena, by virtue of which one communicative form can signal as many as seven different kinds of information¹³, makes suprasegmental communication a difficult matter, since conversationalists not sharing the same background can never be completely sure how to appropriately signal and interpret, encode and decode, the information they wish to exchange.

In the following four chapters the diverse and complex manner in which the three groups of suprasegmental forms - the western (i.e. English and German) and Chinese forms of *intonation*, the *overall features* and voice quality *settings* - contribute to communication will be discussed in detail¹⁴. Each chapter begins with a presentation of the forms to be discussed, following which a description of their most important communicative functions - linguistic, paralinguistic and extralinguistic - is given, based on the evidence in the research literature. Since the focus of this work is on the communication of (interpersonal) attitude, particular attention will be devoted to the manner in which the communicative forms signal this kind of information, both in general terms and in terms of Deborah Tannen's metamessage discussed in Chapter 1.

¹³ In addition to those named here, other functions of these features discussed in the literature include the grammatical, informational and illocutionary functions, the first two subcategories of the linguistic function discussed above (Couper-Kuhlen 1986: 111f.) For a discussion of the indexical function of intonation, see also Crystal 1975: 88ff.

¹⁴ The emphasis in the discussions to follow will be on the signaling of communicative meaning - in terms of linguistic, paralinguistic and extralinguistic information - by the forms of intonation, the overall features and voice quality *settings*. Therefore, the long-term, (quasi-) permanent features mentioned in passing above will be not be included. Of the three kinds of indexical meaning given here, only those relating to speech community membership and speaker personality are of interest to the present work.

2. Forms and functions of English and German intonation

This chapter deals with the signaling of communicative meaning by intonational means in German and English¹⁵. As the first step, the forms of intonation in these two languages, defined according to the tonetic approach to western intonation, are presented briefly¹⁶, following which the communicative functions of these forms are discussed in Section 2.2.

2.1 The communicative forms of English and German intonation

The discussion begins with a presentation of the English and German basic unit of tonality, the intonation-group, and its subdivision into the units *prehead*, *head*, *onset* and *nucleus*. After this we discuss how these units are signaled by length, loudness, pitch and pause.

2.1.1 Units of tonality - the intonation-group

In the description of Western intonation the most basic functional unit is the *intonation-group*, conceivable as the intonational tune, or *pitch-contour*, of a stretch of speech which on the semantic level broadly corresponds to one unit of information¹⁷. On the syntactic plane, intonation-groups generally coincide with either clauses or sentences, a point discussed in more detail below. They are divided into four smaller units, the *prehead*, *head*, *nucleus* and *tail*, the first forming the *prenuclear segment*, *nucleus* and *tail* constituting the *nuclear segment*. Fig. 5 shows an English and a German utterance¹⁸ subdivided into these four units.

Prenuclear segment			Nuclear segment	
Prehead	Head		Nucleus	Tail
1	A	<i>dog_ is a man's best</i>	<u><i>friend</i></u>	
2	A	<i>dog_ is a</i>	<u><i>man's</i></u>	best <i>friend</i>
3	A		<u><i>dog</i></u>	is a man's <i>best friend</i>
1	Ein	<i>bellender Hund</i>	<u><i>beißt</i></u>	nicht ¹⁹
2	Ein	<i>bellender Hund</i>	<u><i>beißt</i></u>	
3	Ein	<i>bellender</i>	<u><i>nicht</i></u>	
			<u><i>Hund</i></u>	beißt <i>nicht</i>

Fig. 5 The units of the English/German intonation-group

As Fig. 5 shows, there is always more than one way of effecting this subdivision, each rendering differing from the others in the position of its *nucleus* relative to the corresponding

¹⁵ Due to the scarcity of tonetically oriented analyses of German intonation, any discussion of the functions of German forms of intonation requires extrapolation from observations on the English language. In the many cases where the meanings of German certain intonation forms are not discussed in the literature, the functions of their English counterparts will therefore be tentatively applied. The formal and functional similarities of English and German intonation which allow this line of action are discussed in Fox 1984, Gibbon 1998, and Grice & Bauman 2000. For discussions of differences in the intonation of the two languages, see Bald 1975, Andersen 1978, Esser 1978, Russ 1981 and Markus 1992. One of the aims of the experimental study in Part Three will be to examine the validity of the functions thus posited for these German intonation forms.

¹⁶ The following discussion is based in the main on the following tonetic works: Pike 1945, O'Connor and Arnold 1973, Crystal 1969, Halliday 1967a, 1970, Cruttenden 1986, Couper-Kuhlen 1986 and Tench 1995 for the English forms of intonation, and von Essen 1964, Pheby 1975, 1984 and Fox 1984 for the German forms. Palmer 1924, Kingdon 1958, Schubiger 1958 and Gimson 1989 were also consulted.

¹⁷ In the early days of intonation research, the intonation-group was thought by some to possess a direct physiological correlate, a unit referred to as the *breath-group*. (See Lieberman 1967, 1980). Later evidence against this hypothesis, presented among others by Ohala & Hirano (1967), led to the rejection of this theory. (See discussion in Couper-Kuhlen 1986: 73 ff.)

¹⁸ The term 'utterance' is used in linguistics to refer to a stretch of speech not (yet) described in terms of phonological or syntactic units (Couper-Kuhlen 1986). Another definition proposed by Crystal (1985) is: 'a stretch of speech preceded and followed by silence or a change of speaker' (1985: 322). The term 'message', adopted from discourse analysis, is used with the same meaning here.

¹⁹ The English example in Fig. 5 is adopted from Tench (1995: 14). The English translation of the German sentence is: 'A dog that barks does not bite'. Unless otherwise indicated, all German examples given in this chapter are direct translations of the English examples they appear in connection with.

utterance. As each individual rendering communicates a different communicative meaning, the location of the *nucleus* relative to the corresponding utterance is of functional, communicative, importance. While its precise location is determined by language-specific rules, the rules of *tonicity*, the general rule in English and German is that the *nucleus*, or *nuclear syllable*, most often falls on the last lexical item of utterances. In each case the *nucleus* marks what is referred to as the *focus* of the utterance, broadly definable as the word containing the most important item of information within the speaker's message²⁰. Coming now to the location of the other three units of the intonation-group, the *tail* extends from the syllable immediately following the *nucleus* up to the end of the utterance, while the *head* begins on the first stressed syllable of the intonation-group, referred to as the *onset*, or *onset syllable*, and ends on the syllable immediately preceding the *nucleus*. Finally, the *prehead* consists of any unstressed syllables preceding the *head*.

In the tonetic approach, the extension of each *intonation-group* relative to the utterance level (a phenomenon called *tonality*), the location of its *nucleus* (*tonicity*) and the pitch-patterns of the nuclear and prenuclear segments, the *nuclear* and *prenuclear pitch-patterns*, constitute the communicative forms of western intonation. The *nuclear pitch-patterns*, also referred to as the *nuclear tones*²¹, are seen as contributing most importantly to intonational meaning, while the pitch-patterns of the prenuclear segment are ascribed a lesser, auxiliary, role²².

It was observed above that on the syntactic plane intonation-groups may coincide with either clauses or sentences. While it is sometimes stated that the most frequent, *unmarked*, association is between the intonation-group and the sentence (Halliday 1975, Tench 1995), studies of English and German spontaneous speech have in fact shown that intonation-groups correspond most frequently to clauses (Crystal 1975, Fox 1984). To take account of this fact, a distinction is generally made between *major* and *minor* intonation-groups: Major intonation-groups are the intonational correlates of whole sentences, while minor intonation-groups coincide with smaller syntactic units, such as (sub- or coordinated) clauses and phrases²³.

Precisely how a given stretch of speech, consisting of any number of utterances, may be divided into major and minor intonation-groups, is an important factor in the determination of intonational meaning. The demarcation of one intonation-group and the placement of its intonation-group boundaries also depend on the correct identification of its internal and external points of delimitation, the *nucleus*, the *onset* and the *intonation-group boundaries* (Cruttenden 1986). The signaling of these points through the action of length, loudness, pitch and pause is therefore briefly discussed as follows.

2.1.2 The onset, the nucleus and the intonation-group boundary

2.1.2.1 The onset

The onset is defined as the first syllable of an utterance marked by what is often referred to as simple, or rhythmic stress, signaled in the main by an increase in duration (*length*) and intensity (*loudness*), sometimes accompanied by rise or fall in *pitch*²⁴.

²⁰ According to the definition suggested by Jackendoff (1972), the focus is 'the information in [a] sentence that is assumed by the speaker not to be shared by himself and the listener' (Couper-Kuhlen 1986: 42). For other definitions of the focus, see Bolinger 1972 b, Chafe 1974, 1976 and Halliday 1967b.

²¹ The terms *tonality*, *tonicity* and *tone* were coined by Halliday (1967b, 1970). The term (nuclear) *tone* is avoided here to avoid confusion with the Chinese (lexical) tones (See Chapter 3).

²² See discussion in Selting 1995: 23f. To be fair, it must be pointed out that this neglect of the prenuclear pitch-patterns is particularly widespread in German tonetic descriptions, but less so in English tonetic literature. This will become evident in the discussion of German intonational meaning below.

²³ The terminology used here in the discussion of syntactic units follows Greenbaum & Quirk 1990.

²⁴ The presence of a rise or fall in pitch leading onto the onset syllable seems to be determined by at least two factors, the pitch of the prehead and the prenuclear pitch-pattern which is thus being initiated. For example, in a high falling head which is preceded by a low prehead there is necessarily a rise in pitch onto the onset.

2.1.2.2 The nucleus

The nucleus is broadly describable as the most prominently stressed syllable of an utterance. In contrast to the rhythmic stress most commonly found on the onset syllable, the prominence of the nuclear syllable is due not only to an increase in length and loudness, but also to a sudden, marked increase or decrease in pitch referred to as a *pitch-obtrusion* or *pitch-accent* (Bolinger 1958)²⁵. Not necessarily the longest, loudest or even the highest syllable within an utterance, the most important feature of the *nucleus* is generally agreed to be this marked change in pitch, which may occur both onto or from the nuclear syllable. The pitch-movement involved here is therefore usually one of the following four possibilities: (1) a *step up to*, (2) a *step down to*, (3) a pitch-movement *down from*, and (4) a pitch-move-ment *up from* the nuclear syllable (Cruttenden 1986: 48). The different positions which can be occupied by the nucleus with respect to the corresponding utterance are referred to as *patterns of tonicity*.

2.1.2.3 The intonation-group boundary

The intonation-group boundary, separating adjacent intonation-groups, may be signaled in a number of different ways, the three most important being: (1) junctural features, such as utterance-final slow-down (lengthening), often followed by a pause²⁶; (2) pitch-declination throughout the utterance; And (3) a marked, clearly perceivable, change in pitch between consecutive utterances (Crystal 1969: 204 ff.).

While the final pause, which may be of varying length²⁷, seems to be more typical of major intonation-group boundaries, utterance-final slow-down is clearly a feature of both major and minor boundaries: From Pike's (1945) discussion, it emerges that final slow-down is present at most minor boundaries, *viz.* those separating co- and subordinate clauses, individual units in lists, and topicalized units, such as adverbial phrases and topics (1945: 31 f. and 45 f.). In the absence of the pause, preboundary lengthening increases to compensate (Crystal 1969: 206). Coming now to the second boundary, *pitch declination*, first observed by Pike (1945) and rediscovered by Pierrehumbert (1979, 1980), may be broadly defined as the phenomenon that the pitch of a speaker's voice is generally lower at the end of an utterance than at the beginning²⁸. While declination is generally agreed to originate in the decrease in subglottal pressure during speech, researchers disagree as to whether this phenomenon is purely phonetic or consciously controllable, in which case it could serve as a communicative, phonological, phenomenon in its own right²⁹. Indeed, the results of some studies indicate that declination may not only be controlled but also suspended by certain laryngeal control mechanisms (Cohen et al. 1982, Ladd 1984)³⁰. It also appears that *top-line declination*, affecting the pitch of high-pitched, stressed syllables of utterance, is for the most part dependent on physiological factors and outside conscious control, while *bottom-line declination*, which affects the pitch of low, unstressed syllables, may be actively controlled to a certain degree. As Cruttenden (1986) notes, there are also indications that bottom-line declination plays a certain, though indirect, role in the signaling of intonation-group boundaries: These are often marked by the suspension of declination, in whose place there is a sudden, relatively pro-

²⁵ According to Crystal (1969), Couper-Kuhlen (1986) and Cruttenden (1986), sudden rises and falls in pitch very similar to *pitch-accents* may also be found on onset syllables and within the *head*.

²⁶ While many tonetic descriptions of western intonation systems emphasize the role of the pause, the most frequently encountered junctural signal is in fact not the pause, but utterance-final, or preboundary, slow-down, resulting from a lengthening of the last two, maximally three syllables of an utterance (Pike 1945). (See also Cruttenden 1986, Trim 1959/1973 and Pike 1945).

²⁷ There appears to be no agreement on the relative length of these final, phonological, pauses as compared to the length of non-phonological, hesitation pauses which occur in non-final position. (See Brown et al. 1980 and Cruttenden 1986: 39)

²⁸ Other studies of this phenomenon include Ohala 1978, Thorsen 1980, Cohen et al. 1982 and Ladd 1984. See also the overview provided by Hirst & Di Cristo (1998).

²⁹ In view of this uncertain status, Connel & Ladd (1990) have suggested the generic term of *downtrend*. It seems that the terms *downtrend* and *downstep* often found in the literature relate to phonological manifestations of *downtrend*, while *declination* is used for strictly phonetic manifestations of this phenomenon.

³⁰ See Cohen et al. 1982 and Ladd 1984. As Cohen et al. observe, the calculation of declination is generally based on that of the *bottom-line*, which is regarded as primary (1982: 261). The results of their calculations for English, Dutch and Japanese indicate that, at least in these languages, shorter utterances contain a comparatively stronger degree of declination than longer ones (1982: 259).

nounced shift in pitch or a change of direction extending over the last unaccented syllables (1986:126)³¹. This phenomenon corresponds to the third boundary marker named above.

2.1.3 The pitch-patterns in the prenuclear and nuclear segments

2.1.3.1 The prenuclear pitch-patterns (heads)³²

The prenuclear segment was described above as consisting of the *head* and *prehead*. While the *head* is defined as containing any and all the rhythmically stressed syllables preceding the *nucleus*, the *prehead* consists of all the unstressed syllables preceding the *head* (or, in the absence of a *head*, the *nucleus*). In English and German, the *prehead* generally contains only unstressed syllables, whose pitch is largely determined by that of the following *onset*. Its contribution to intonational meaning is therefore generally agreed to be minimal, a position which is adopted here. Discussions of prenuclear pitch-patterns in this work will therefore be mostly limited to those of the *head*³³.

The number of prenuclear pitch-patterns often varies between individual tonetic descriptions of English intonation. Broadly speaking, however, three major types of prenuclear pitch-movement are recognized: *level*, *rising* and *falling heads*. *Level heads*, including by definition gently descending or ascending pitch-movements, are further differentiated according to their overall pitch-height into *high*, *mid* and *low* types, and *rising* and *falling heads* are subdivided into *simple*, *stepping* and *glissando* types, the latter distinction pertaining to the pitch-movements of the unstressed syllables: Whereas in simple falling or rising *heads* the overall pitch-movement - rising or falling - affects all syllables in equal manner, by contrast in the *stepping* and *glissando* types the pitch-movement is limited to the stressed syllables. While in *stepping heads* unstressed syllables adopt the pitch-height of the stressed syllables immediately preceding them (as in the *stepping head*), in *glissando heads* they form an intermediary *falling* or *rising* pitch-sequence which effectively continues the pitch-movement traced by the preceding stressed syllable³⁴. The difference between *simple*, *stepping* and *glissando* falling *heads* is shown in Fig. 6 below.



Fig. 6 The *stepping*, *glissando* and simple types of falling *heads*³⁵

Coming now to the German prenuclear pitch-patterns, the earliest description by von Essen (1964) recognizes two major kinds of *head* - *low* and *mid* (either level or gently descending)

³¹ As Cruttenden also observes, however, such preboundary *down-* and *upshifts* in pitch, the latter also referred to as *resets* do not always occur here (1986: 41), as fact which also applies to the suspension of declination described here. The same degree of declination has been found to apply across two or more intonation-group boundaries, and may affect a number of the intonation-groups contained in one *paratone*, the intonational equivalent of the paragraph (See Couper-Kuhlen 1986:189 f.). According to Couper-Kuhlen, *pitch resets* are often - but again not always - found between consecutive *paratones* (1986: 189 f.).

³² In the tonetic literature prenuclear pitch-patterns are generally referred to as *heads*, the only exception being Kingdon 1958, where this term relates to the *onset*.

³³ One of the few linguists to discuss English *preheads* in more detail is Crystal (1969), who distinguishes between five different pitch-heights, ranging between 'extra-high' and 'extra-low' (1969: 233).

³⁴ Tench (1995: 132f.). See also O'Connor & Arnold 1973, where the *stepping* and *glissando heads* are included under the headings 'stepping' and 'sliding' *heads*.

³⁵ Large dots represent stressed syllables, small dots correspond to unstressed syllables.

heads (1964:23f.). This number is raised slightly in Pheby (1975) and Fox (1984)³⁶: Working with the same three major types as discussed for the English system, Pheby distinguishes between the following four major types: the *level head*, the simple *falling head*, the ‘stepping’ variety of the *falling head* and the ‘stepping’ *rising head* (1975: 62f). Fox on the other hand differentiates between three prenuclear pitch-patterns - high, low *level* and low *rising*, the pitch-movement of the latter referred to as ‘undulating’ (1984: 30f.)³⁷. Interestingly, no mention is found in these two analyses of *glissando* and *stepping heads* in German.

2.1.3.2 The nuclear pitch-patterns

Beginning with the pitch-accent, leading either to or from the *nuclear* syllable, the nuclear segment extends throughout the *tail* up to the end of the intonation-group. In most tonetic analyses, different types of nuclear pitch-patterns are described with two sets of criteria - *nuclear direction* and *nuclear pitch-range*. A third criterion, *tail-slope configuration*, is found in Couper-Kuhlen (1986).

(1) Nuclear direction

This criterion, often referred to as *primary*, relates to the different directions which may be traced by nuclear pitch-patterns. The five possible nuclear directions in English and German are the simple *fall* and *rise*, the bidirectional *fall-rise* and *rise-fall*, and the *level* direction³⁸. While no consensus has been reached with regards to the existence of *level* nuclear pitch-patterns in English³⁹, the German intonation system is generally agreed to consist of five distinct nuclear directions: *falling*, *rising* and *level*, *fall-rising* and *rise-falling*⁴⁰.

(2) Nuclear pitch-range

The secondary criterion of *nuclear pitch-range* - also referred to as *accent-range* - is generally described in connection with, and as a function of, the *pitch-height* of the nuclear syllable. Most tonetic descriptions of English and German intonation distinguish between three nuclear pitch-heights, *high*, *mid* and *low*, and three pitch-ranges, *wide*, *mid* and *narrow*. By convention, the pitch-height of a falling pattern relates to the starting-point of its descending pitch-movement, while the pitch-height of a rising pattern corresponds to the highest pitch attained by its ascending pitch-movement. As in most tonetic analyses all rising types end high and all falling nuclear patterns end low, the number of types of nuclear pitch-patterns distinguishable with the criterion of *pitch height/range* is relatively limited: Maximally three (low, mid and high) *falls* and *rise-falls*, three *rises* and *fall-rises* and, if allowed for, three *levels*. In fact, most tonetic models provide for only one *rise-fall* and one *fall-rise*, starting high and low, respectively⁴¹. The only exception in English is Pike (1945), who not only distinguishes between four pitch-heights, but also provides for a larger number of possible starting and ending points - and thus a larger number of possible accent-ranges - per nuclear direction. More will be said of Pike’s model below.

Coming now to the German nuclear pitch-patterns, the same overall methods of description are found here. However, the number of nuclear types varies from one description to the next. Thus, in addition to mid, high and low *falls*, *rises* and *levels*, Pheby (1975) differentiates between mid, high and low *fall-rises*, but allows for only one *rise-fall*, starting at low-mid

³⁶ There is no discussion of prenuclear pitch-patterns in Pheby 1984.

³⁷ As one minor difference between German and English intonation patterns, a slight upward trend appears to be characteristic of unstressed syllables in German ‘stepping’ *heads*, a phenomenon described by Fox in terms of a gentle ‘undulating’ movement (1984:34).

³⁸ The *rise-fall* is regarded by some analysts, as for instance Tench (1995), as a variant of the *fall*.

³⁹ While Pike 1945, Crystal 1969, Cruttenden 1986 and Couper-Kuhlen 1986 all allow for level English nuclear pitch-patterns, these are not recognized by Halliday (1967a, 1970) or Tench (1990). Here, levels are seen as ‘rare variants’ of the simple rise.

⁴⁰ While Pheby’s descriptions (1975, 1984) include all five nuclear tones, Fox (1984) sees the *fall-rise* as a variant of the simple *rise*.

⁴¹ One exception here is Tench (1995), who differentiates between three *fall-rises* and two *rise-falls*. See below.

pitch (1975: 66). In contrast to this, Fox (1984) has high, mid and low *rises*, *rise-falls* and *fall-rises*, but recognizes only low and mid *rises* and *levels* (1984: 22f.).⁴²

As a final point, most tonetic descriptions of English intonation also provide for the occurrence of more than one nuclear pitch-pattern within one intonation-group, a phenomenon referred to as *compound nuclear pitch-patterns*. Again, both the types and numbers vary stronger between individual descriptions. Thus, Halliday (1970) differentiates between two compound nuclear tones, the *falling* + *low rising* and the *rise-falling* + *low rising* (1970: 9), while Couper-Kuhlen (1986) suggests four basic tone sequences, the *fall* + *rise*, the *rise* + *fall*, the *rise-fall* + *rise* and the *rise-fall* + *fall-rise* (1986: 98 f.), to which other minor types are added.⁴³

(3) Tail-slope configuration

The tail was defined above as consisting of all the unstressed and (rhythmically) stressed syllables following immediately after the nuclear syllable. Though different tail configurations have long been recognized, most tonetic descriptions see the pitch-movement within this part of the intonation contour as a continuation of the preceding nuclear pitch-pattern, and thus as a purely phonetic phenomenon devoid of communicative function⁴⁴.

As Couper-Kuhlen (1986) has shown, however, one and the same nuclear pitch-pattern may in fact occur with different tail-slope configurations, which also signal subtle differences in meaning. Therefore, the conditioning relationship between type of nuclear tone and configuration of the tail-slope may not be quite as strong as formerly believed. Couper-Kuhlen proposes a three-way distinction, between (1) *gradually* descending/ascending tail-slopes, also called *gentle tail-slopes*, or *tails*, (2) sharply descending/ascending *steep tails* and (3) slope-less tails, also referred to as *flat tails*, which level out the nuclear movement (1986: 86).

In cases where a *flat tail* is preceded by a high or mid falling *nucleus* in turn preceded by a mid/high *head*, the nuclear type, which thus contains a sudden and abrupt step-down in pitch towards the nuclear syllable is referred to here as a ‘*drop*’ *nucleus*. Fig. 7 shows the four types of tail-slope configuration differentiated here.

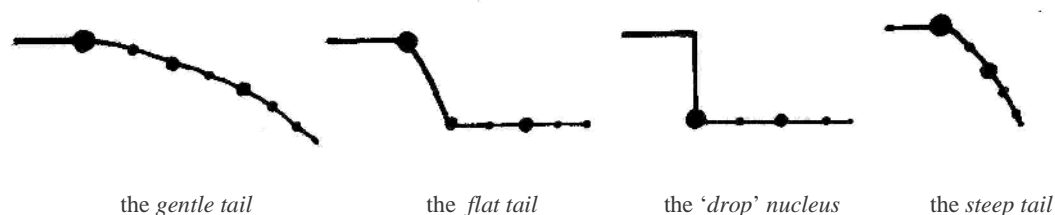


Fig. 7 Four types of tail-slope

A distinction of this kind, between different configurations of tail-slope is not found in the German tonetic literature.

⁴² Though von Essen (1964) does not explicitly differentiate between different variants of nuclear pitch-patterns, the data he discusses clearly include high, low and mid *falls*, mid and low *rises*, but only one (mid) *level* tone. There are, however, no *fall-rises* or *rise-falls* in von Essen’s description.

⁴³ Interestingly, no compound nuclear tones are described in the tonetic descriptions of German intonation under discussion here.

⁴⁴ In effect, this finding by Couper-Kuhlen is in opposition to the general view that, in contrast to *rises*, which extend over the entire nuclear segment, in English *falls* the falling pitch-movement is always completed no later than the syllable immediately following the *nucleus* (O’Connor & Arnold 1973, Halliday 1970).

2.2 The communicative functions of English and German intonation

The communicative functions of the English and German forms of intonation which were defined above are presented in the following order: After a brief discussion of *tonality* and *tonicity*, the functions of *nuclear direction* and *pitch-height/range* and *tail-slope configuration* are discussed in detail, followed by a brief account of the most important functions of the *pre-nuclear pitch-patterns*. As mentioned above, the focus here will be on three types communicative functions: (1) *linguistic*, pertaining to discourse cohesion, the internal, informational, structure of the speaker's message⁴⁵, (2) *paralinguistic*, relating to his emotional and attitudinal disposition at the time of speaking, and (3) *extralinguistic*, relating to his speech community membership and certain - genuine or judged - traits of his personality.

2.2.1 The functions of tonality and tonicity

While *tonality* and *tonicity* can also contribute to the signaling of para- and extralinguistic information, it seems that their primary functions are of a linguistic nature, serving the establishment of discourse cohesion. As this involves communicating the information structure of the speaker's message, the linguistic functions of *tonality* and *tonicity* are sometimes also referred to as their 'informational' functions (Couper-Kuhlen 1986, Tench 1995). *Tonality* and *tonicity* have also been found to accompany certain syntactic structures. These so-called 'grammatical' functions will also be included very briefly.

2.2.1.1 Discourse cohesion - the informational functions of tonality and tonicity

The signaling of discourse cohesion was described in Chapter 1 as the first step towards the establishment of conversational involvement. Its objective is to ascertain that the information contained in the speaker's message is interpreted correctly and in the right order. The two most important steps in the signalling of discourse cohesion involve, first, presenting, or 'packaging' (Chafe 1986), one's message in portions which can be easily processed by the listener and, second, labelling the most important item of each portion. Consequently, the informational function of *tonality* is the subdivision of speaker messages into smaller portions, or *units of information* (Halliday 1967a), while *tonicity* marks the main item, or *focus*, of each unit of information.

Tonality - the packaging of information

Tonality is thus broadly defined as the process of subdividing a message into a sequence of intonation-groups, each of which corresponds to one unit of information. While on the one hand this process is regulated by *tonality* rules⁴⁶, on the other the subdivision of a message into individual units also depends on how the speaker chooses to present the information contained therein. Consequently, there is always more than one way of going about such a subdivision, a fact which already pointed out at the beginning of this section. In the following examples, an English and a German message, each consisting of one utterance, are divided into intonation-groups in three different ways⁴⁷.

- | | |
|---|---|
| (1) John saw the <i>play</i> yesterday
 <i>John</i> saw the <i>play</i> yesterday
 <i>John</i> saw the play yesterday | (2) Peter ist gestern ins <i>Theater</i> gegangen
 <i>Peter</i> ist gestern ins <i>Theater</i> gegangen
 <i>Gestern</i> ist Peter ins <i>Theater</i> gegangen |
|---|---|

⁴⁵ As discussed in Chapter 1, the signaling of discourse cohesion involves the communication of the internal structure of the speaker's message, which may contain any number of sentences.

⁴⁶ Detailed discussions of the English and German rules of *tonality* are found in Cruttenden 1986 and Tench 1995 and in Pheby 1975 and 1984, respectively.

⁴⁷ The English example is taken from Couper-Kuhlen (1986: 122). The English translation of the German example is: 'Peter went to the theatre yesterday.'

Tonicity - the marking of the focus

Tonicity is broadly defined as the marking of the *nucleus* of a given intonation-group, this in turn reflecting the most important item, or *focus*, of the corresponding utterance. Since one intonation-group normally contains one, maximally two *nuclei*, the number of *nuclei* per utterance is also determined to some degree by the number of intonation-groups it is divided into. The precise location of the *nucleus* within a given intonation-group, is regulated by the relatively complex, language-specific rules of *tonicity*⁴⁸. In English and German *tonicity* the following three factors appear to be most important for the location of the *nucleus*: (1) The type of *focus*, i.e. *narrow* or *broad*, which is to be signaled; (2) the *word type* which is involved in *broad focus*; and (3) the *syntactic rules* which are activated. In *narrow focus*, where one particular item is being highlighted, the general rule for English and German is that the *nucleus* always falls on this very item. *Narrow focus* may involve the signaling of a contrast, as in example (3) below, or the highlighted item is seen as new or unexpected to the listener - either by comparison to other items in the utterance or within the communicative context provided, as in example (4). *Broad focus*, on the other hand, is defined as a condition in which the speaker's message in its entirety is signaled as new, an example for which is shown below in (5). As example (5) illustrates, the general rule here is that the *nucleus* falls on the last lexical item of an utterance, this excluding by definition all items which are regarded as non-lexical in the respective language. While in English this involves the relatively limited group of grammatical morphemes, in German the list of word-types regarded as 'non-lexical' in this context is much longer, including for example, the prefixes of separable words and often even verbforms, as is shown in example (6). Finally, in both languages the issue of *broad focus* is further complicated by syntactic rules which also regulate the position of the *nucleus*. In English, this involves for example what Cruttenden (1986) calls 'event sentences', which typically contain verbs relating to (dis-) appearance and misfortune and where the *nucleus* usually falls on the subject⁴⁹. *Nucleus* position in event sentences is illustrated in example (7).

(3) *Contrastive focus*: She's not my girlfriend anymore. She's my *ex*-girlfriend⁵⁰.
Sie ist nicht mehr meine *Freundin*. Sie ist meine *Ex*-freundin.

(4) *Narrow focus*: [Q: What will you have to drink?]
(non-contrastive) A : I think I'll have a cup of *tea*. Ich glaube, ich nehme einen *Tee*.

(5) *Broad focus*: [Sitting on the balcony at night]:
Look, there's a falling *star*! Guck mal, da ist ja eine *Sternschnuppe*!

(6) Event sentence: [Watch out!] The *chimney*'s falling down!⁵¹
[Pass auf!] Der *Kamin* fällt herunter!

2.2.1.2 The grammatical functions of tonality and tonicity

While going about their informational functions, *tonality* and *tonicity* come into contact with the syntactic structures of utterances. It has been found that certain patterns of *tonality* and *tonicity* regularly and typically co-occur with particular syntactic structures. In some cases, it even seems that their very presence can serve to contrast, or differentiate between, such structures. Examples (8) and (9) show English utterances in which *tonality* and *tonicity* would seem to fulfill such contrastive functions⁵²:

⁴⁸ A detailed discussion of the English rules of *tonicity* is found in Cruttenden 1986 and Tench 1995. The German rules of *tonicity* are presented in Pheby 1975 and 1984. (See also Fox 1984: 53 ff.).

⁴⁹ See Cruttenden 1986: 83.

⁵⁰ The English example is adapted from Tench 1995: 62.

⁵¹ The English example is taken from Cruttenden 1986:83.

⁵² Note that in many of the following cases *tonality* is in fact supported in this function by the nuclear pitch-patterns, which in their informational function (see below) serve to indicate finality vs. non-finality. Therefore, the so-called 'grammatical' function of tonality - and tonicity too - is in doubt. (See Barry 1981, Couper-Kuhlen 1986: 151 ff.).

(8) *Tonality*: Defining vs. non-defining appositions and relative clauses⁵³:

- (a) | They sent *Joan* | a waitress from the *hotel* || vs. | They sent *Joan* | a waitress | from the *hotel* ||
(b) | Tom *Jones* | the *singer* | comes from *Wales* || vs. | Tom Jones the *singer* | comes from *Wales* ||

(9) *Tonicity*: Comment adjuncts vs. adjuncts of manner⁵⁴:

- | They didn't come *happily* || vs. | They didn't come | *happily* ||

On the other hand, in examples (10) and (11) there is a looser relationship between *tonality* and *tonicity* and syntax⁵⁵:

(10) *Tonality*: The marking of direct speech interspersed by a comment clause:

- | *What* | people *ask* | do you *make* of him? ||
| *Was* | fragten die *Leute* | habe er hier zu *suchen*? ||⁵⁶

(11) *Tonicity*: The marking of final vocatives, included in the tail:

- | I didn't say *that* Mr. *Powel* ||
| Das habe ich nicht so *gemeint* Peter ||

2.2.1.3 The para- and extralinguistic functions of *tonality* and *tonicity*

The role of *tonality* and *tonicity* in the communication of para- and extralinguistic information has received only very little scholarly attention and the little research that has been done is almost entirely limited to the English language. While the role of *tonicity* in the signaling of paralinguistic information is still in doubt, there is a fair amount of evidence that *tonality* has important contributions to make here. Halliday (1970) observes that *tonality* patterns in English often reflect the formality of the speech situation: In formal situations, where speech is relatively slow and utterances are long and full of information, these tend to be divided into a larger number of intonation-groups than in informal speech (1970: 4). Gumperz (1982a) has pointed out that among the two *tonality* patterns shown here in example (13) version (b), containing three intonation-groups, conveys a higher degree of emotional intensity than version (a), containing only one intonation-group. Example (14) suggests that this may also be the case in German.

- (13) (a) | I said sit *down* || vs. (b) | I *said* | sit | *down* || (Gumperz 198a: 110)
(14) (a) | Ich *sagte* |: | Setz dich *hin* || vs. (b) | Ich *sagte* |: | Setz | dich | *hin* ||

With respect to the possible extralinguistic functions of *tonality* and *tonicity*, studies in comparative intonation have shown that the non-standard patterns of *tonality* and *tonicity* found in the English speech of non-native speakers invariably reflect the *tonality/tonicity* patterns of their mother tongue and may thus serve to signal speech community membership. Vanderslice et al. (1972) and Gumperz (1982) have described non-standard patterns of *tonality* and *tonicity* in Hawaiian English (a Creole) and Indian English. Example (15), taken from Gumperz, Aulakh & Kaltmann 1982, gives an impression of the *tonality* patterns which may be found in Indian English⁵⁷:

- (15) | So *therefore* | *normally* | most of the young *people* | who *are* | getting *married* | *newly* |
they have got no *families* ||

⁵³ These examples are adopted from Tench 1995: 42

⁵⁴ Example from Tench 1995: 71

⁵⁵ The English examples are from Couper-Kuhlen 1986: 147.

⁵⁶ The English translation of this sentence is: 'What, asked the people, was his business in coming here'

⁵⁷ Gumperz 1982a: 24. See also the discussion in Gumperz 1982b: 121 f. and Cruttenden 1986: 143. Interestingly, very similar non-standard *tonality* and *tonicity* patterns are found in English and German utterances spoken by native speakers of Chinese.

2.2.2 The functions of nuclear direction

2.2.2.1 Discourse cohesion - the linguistic functions of nuclear direction⁵⁸

According to Tench (1995), the linguistic function of nuclear direction in the service of discourse cohesion may be described as the signaling of the ‘information status’ of the speaker’s message, in terms of *completeness/finality* and *non-completeness/non-finality*. In their basic linguistic function *falling* nuclear pitch-patterns signal that a message is complete, while *rises* imply that it is incomplete and thus awaiting, or expecting, completion, either by the speaker or the listener (1995: 80). Regarding the basic informational meanings of the bidirectional nuclear directions, not included as such in Tench’s discussion, it is suggested here that *rise-falls* and *fall-rises* may serve as intensified forms of the *falls* and *rises*, serving to emphasize and call special attention to the completeness or incompleteness of a given message⁵⁹.

As was also observed for *tonality* and *tonicity*, in their function as cohesion markers the nuclear pitch-patterns also come into contact with the syntactic structures of utterances. Therefore, the linguistic information which is signaled by a nuclear tone also depends to a certain degree on its position within the speaker’s message - in terms of *final* or *non-final* in a sequence of clauses and sentences (1995: 80f.). Thus, while the general rule is that *falling* nuclear pitch-patterns signal completeness and finality, while *rises* signal the message of incompleteness and non-finality, a further differentiation of meaning is necessary for rises: Whereas a *rise* in non-final position generally signals that the speaker is about to continue, a *rise* in final position does not specify whether it is the speaker or the listener who is expected to complete the message.

Coming now to the *rise-fall* and the *fall-rise*, Tench sees the *rise-fall* as an intensive version of the simple *fall*. Therefore, it seems possible to associate it with a certain degree of ‘strength’, emphasizing the completeness and finality of the speaker’s message. Finally, the *fall-rise*, like the simple *rise*, is associated with two slightly different meanings: While in non-final position it ‘highlights the message content’, in final position the *fall-rise* communicates ‘implication’ (1995: 84).

Turning now to the functions of nuclear direction in German, a basic association between *falls* and *completeness* or *finality* and between *rises* and *incompleteness* or *non-finality* is also found in the German tonetic literature. Thus, Fox (1984) suggests the very similar labels of ‘closedness’ for *falls* and ‘openness’ for *rises* (1984: 60). While his framework excludes *fall-rises* and the meaning he proposes for *rise-falls* (‘assertion’) is of an interpersonal nature, Fox does associate *levels* in non-final position with a feeling of ‘incompleteness’⁶⁰. This suggests that the above discussed English linguistic meanings for *falls* and *rises* are also applicable to their German counterparts. Unfortunately, nothing concrete may be said about the functions of *rise-falls* and *fall-rises*, as the meanings of these pitch-patterns are not discussed in the tonetic literature. However, given the structural and functional similarities between English and German, to which attention was called in the General Introduction, it is conceivable that the English meanings discussed above for these nuclear patterns also apply to their German correlates. The linguistic functions meanings of the four English and the five German nuclear directions are shown in Fig. 8.

2.2.2.2 The paralinguistic functions of nuclear direction

In their paralinguistic function, the English and German nuclear directions serve to communicate speaker attitude, definable in broad terms as the ‘stance’, or ‘position’ he assumes with

⁵⁸ This discussion is based on the tonetic descriptions of English and German intonation named in Section 2.1.2.2, excluding Pheby 1975 and 1984. The Hallidayan approach adopted by Pheby, which distinguishes ‘unmarked’ and ‘marked’ co-occurrences between nuclear pitch-patterns and syntactic forms, is different from the approach followed here.

⁵⁹ Note that these meanings are highly compatible with the meanings Tench ultimately suggests for the bidirectional pitch-patterns, to be included below.

⁶⁰ Von Essen (1964: 15) also associates the (non-final) level pitch-pattern with ‘incompleteness’ and ‘progreddience’. Interestingly, Tench also associates the *level* pitch-patterns, which he regards as ‘rare variants’ of simple *rises*, with ‘incompleteness’ (1995: 81), a fact that illustrates the functional similarity of English and German forms of intonation.

<i>Nuclear direction</i>	<i>Position within message</i>	<i>Linguistic/informational functions</i>
Simple fall	final and non-final	completeness/finality
Simple rise	final	incomplete, to be completed (by speaker or listener)
	non-final	incomplete, to be completed (by speaker)
Rise-fall	final and non-final	emphasis of completeness
Fall-rise	final	highlighting
	non-final	implication
Level	non-final	incompleteness
	final	-

Fig. 8 The linguistic functions of nuclear direction in English and German

regards to an object, an event or a person. While this first definition may sound quite simple and straightforward, it is important to note that the signaling of attitude in suprasegmental communication, particularly in intonation, actually involves two messages of a rather different nature: (1) the speaker's attitude towards the message he is imparting, and (2) the (interpersonal) attitude towards his interaction partner. This differentiation will be maintained in the following discussion of the attitudinal meanings of the nuclear pitch-patterns.

Speaker attitude towards the message

The speaker's attitude towards his message is conceivable as the manner in which he presents its content to the listener, i.e. either gently/mildly or strongly/forcefully. In so doing, he signals the degree to which he believes the content of his message to be true⁶¹. According to Halliday (1970), therefore, in its most basic form the speaker's attitude towards his message, as signaled by nuclear direction, may be described in terms of his *certainty* or *uncertainty* with respect to its content⁶². As he writes:

Basically, a falling contour means certainty and a rising contour means uncertainty. This is true of many languages, though by no means all. In English, it takes this particular form: A falling contour means certainty with regard to yes or no. We go down when we know something is positive or negative, we go up when we do not know [...]. (1970: 23)

Thus, while *falls* reflect certainty, sureness and therefore also a certain degree of self-confidence on the part of the speaker, *rising* nuclear patterns indicate that the speaker harbours a certain degree of doubt/uncertainty with respect to (some aspect of) his message. With respect to the attitude expressed by *fall-rises* and *rise-falls*, Halliday suggests that the bidirectional pitch-patterns be regarded as containing '[both] components of meaning with 'a change of mind in the middle': Thus, a *fall-rise* signals that 'it may seem as though all is clear, but in fact there is more involved', while the *rise-fall* conveys that 'there may seem to be a doubt, but in fact all is certain' (1970: 23). In consequence, Halliday sees utterances carrying *fall-rises* as signaling reservation, implying a 'but', whereas *rise-falls* give utterances a strong, assertive feeling (1970: 24). Thus, to Halliday, the attitudinal meanings of nuclear direction are certainty (*falls*), assertiveness (*rise-falls*), uncertainty (*rises*) and reservation (*fall-rises*).

Speaker attitude towards the listener (interpersonal attitude)

The speaker's attitude towards the listener, his interpersonal attitude, was broadly defined as a *stance*, or *position*, he or she adopts with respect to his interaction partner. According to the

⁶¹ What is referred to here as the speaker's attitude towards his message, as signaled by the forms of intonation, should not be confused with the concept of *modality*, often described in similar terms in the literature.

⁶² Note that the degree of this *certainty* is reflected by what is often referred to the speaker's 'degree of emotional involvement' and the utterance's 'emotional intensity'.

definition presented in Chapter 1, interpersonal attitude relates on the one hand to what the speaker sees to be his relative - social or interactional - *standing*, or *power position*, with respect to the listener (*footing*), and on the other to the degree to which he is prepared to identify or associate himself with the listener, his qualities, wants and needs (*affinity*).

Before we come to the manner in which interpersonal attitude is communicated, it should be pointed out that the communication of the speaker's attitude towards his message and that of his attitude towards his interaction partner are in fact not two separate signaling processes but intrinsically linked. This connection is reflected in the fact that both kinds of attitude are signaled by the same intonational means. To illustrate this, we briefly return to Gumperz' and Tannen's research in interpersonal communication. A point to which Gumperz and Tannen repeatedly call attention is the fact that the manner in which a message is presented always has a certain attitudinal, in the sense of interpersonal, effect on the listener. This in effect means that while on the one hand *falls* signal certainty with respect to message content, on the other hand they may also be used to reflect the speaker's self-confidence. This certainty and self-confidence, however, can be interpreted in terms of an assertive or superior standing, or *footing*, towards the listener. In the same manner, while the basic attitudinal meaning of the *rise* is uncertainty with respect to one's message content, a *rise* can also be used - or interpreted - to communicate an interpersonal message of unsureness, lack of self-confidence and thus, inferiority. Thus, the very speech behavior adopted by a speaker who is certain of the content of his message - the falling nuclear pitch-patterns which he adopts to signal this certainty - can at the same time convey an authoritative, superior and dominant interpersonal attitude. Likewise, the *rise* which a speaker uses to communicate uncertainty and unsureness may also be employed or interpreted to signal an inferior attitude, or footing, with respect to his interaction partner.

Returning now to the ongoing discussion of the interpersonal meanings of the nuclear pitch-patterns, Tench (1995), whose description of the attitudinal meanings of the English *falls* and *rises* builds upon this link between certainty/dominance and uncertainty/submissiveness, therefore defines *falls* as signaling the (interpersonal) attitudes of certainty and dominance, while *rises* communicate a submissive or deferent attitude towards the interaction partner:

'A fall indicates the speaker's certainty or dominance in respect of the knowledge, authority and feelings, [while] a rise indicates the speaker's uncertainty and/or deference to the knowledge [...], the authority and the feelings of the person addressed' (Tench 1995: 89).

Coming to *rise-falls* and *fall-rises*, though these do not receive attitudinal meanings derived from certainty and uncertainty in Tench's model⁶³, since the *rise-fall* was defined above as a kind of intensified version of the simple *fall*, it seems possible to associate this pitch-pattern with the emphasis of the speaker's certainty, this in turn giving rise to a heightened degree of forcefulness and dominance. In a similar way, it seems possible to give the *fall-rise* the basic interpersonal meaning of emphasizing the speaker's uncertainty and his deference to the knowledge of the listener, this in turn linkable to a strongly inferior interpersonal attitude, or footing⁶⁴.

Turning now to the interpersonal meanings of nuclear direction in German, no mention of the terms certainty and uncertainty is to be found in the literature. On the other hand, the interpersonal meanings employed here do go in the same direction. Thus, Fox (1984) suggests the basic meanings of 'assertion' and 'appeal' for *falls* and *rises*, the term 'appeal' implying that the listener is being asked to respond to the speaker's message (1984: 59). Interestingly, a little further on, Fox observes that 'assertion' is in fact more appropriate for the *rise-fall* and that he prefers to associate the simple *fall* with the feeling that the speaker's utterance is 'self-sufficient' and 'independent', requiring no response from the listener apart from an acknow-

⁶³ Tench's model does not provide for any attitudinal meanings of neutral, i.e. mid, *fall-rises* and *rise-falls*.

⁶⁴ The validity of this postulation will be examined in Part Three (Chapter 9).

ledge (p.67). Thus, Fox's meanings of 'self-sufficiency' and 'appeal' are highly compatible with the meanings 'certainty' and 'uncertainty' discussed above for the English *falls* and *rises*, and his view of the *rise-fall* as an intensified version of the simple *fall*, signaling 'assertion', also clearly goes in the same direction as the above suggested association of English *rise-falls* with 'forcefulness'. Finally, with respect to utterance-final *levels*, described as most common in ritual expressions, such as greetings and farewells, Fox notes that in normal conversational utterances these are used to communicate knowledge which the listener is thought or expected to be familiar with, so that this pitch-pattern may be said to signal a certain 'triviality' and 'self-explanatoriness' of the message (p. 69). On the interpersonal level this corresponds to an attitude of disinterest and boredom. The attitudinal functions discussed here for the four English and five German nuclear directions are shown in Fig. 9 below:

<i>Nuclear direction</i>	<i>Attitudinal functions</i>	
	<i>towards the message</i>	<i>towards the listener</i>
Falls	certainty, sureness	confidence, dominance
Rises	uncertainty, unsureness	lack of confidence, deference
Rise-falls	strong certainty	strong dominance, assertiveness
Fall-rises	strong uncertainty, reservation	strong uncertainty, deference
Levels (final)	self-explanatoriness, triviality	boredom

Fig. 9 The attitudinal functions of nuclear direction in English and German⁶⁵

2.2.2.3 The extralinguistic functions of nuclear direction

Research into the signaling of speech community membership and speaker personality by intonational means has been almost entirely limited to the English language and only a few remarks on German can be included.

The signaling of speech community membership

With respect to nuclear direction, the most frequently discussed differences between social and regional varieties of English and German concern the use of simple *falls* and *rises*. Thus, in a large number of English dialects, especially those spoken in the Northern and Western areas of Great Britain, *rises* are often used in situations where *falls* would be expected in the RP standard. According to Cruttenden (1986), who gives a brief review of the research in this field, this extensive use of *rises*, particularly conspicuous in statements and sentence-final positions, encountered in cities like Birmingham, Liverpool, Glasgow, Belfast and Tyneside, as well as in Welsh English, is due to the strong Celtic influence in these areas (1986: 139)⁶⁶. Similar non-standard patterns concerning the use of *rises* have also been described for other parts of the English speaking world, such as America, Canada and Australia. Cruttenden (1986) and Bolinger (1989) both call attention to the higher incidence of *rises* in statements in Standard American, a feature which, according to Cruttenden, is especially common in narrative monologues (1986:142)⁶⁷, and the widespread presence of *rises* also appears to be characteristic of Black English Vernacular (Loman 1975, Tarone 1976, Gumperz 1982a)⁶⁸. Finally, in Australia the growing use of rises in statements in the speech of women, adoles-

⁶⁵ Note that in this table the positions of the nuclear patterns within a message are not taken into consideration. This is in accordance with the general view in the literature that in contrast to the linguistic meanings discussed above, which may differ depending on position, speaker attitude is communicated mainly, if not exclusively, by nuclear pitch-patterns in utterance-final position. (See Tench 1995: 87). For this reason the meanings of non-final *levels*, exclusively of an attitudinal nature, are also not included in Fig. 8.

⁶⁶ For studies of Scottish and Irish intonation, see Kenworthy 1978, Currie 1979, Jarmen & Cruttenden 1976 and Pellowe & Jones (1978. See also the discussion in Bolinger 1989: 26 ff.

⁶⁷ Bolinger (1989) includes a detailed comparative discussion of British and American English intonation.

⁶⁸ Black English Vernacular, also referred to as Afro-American, is the speech of the black population of American urban communities. (Crystal 1967: 35) One example for the use of *rises* in statements was discussed above in Chapter 1 ('I don't wanna read').

cents and members of the working-class is being observed since the early 1980s (Horvath 1985, Guy et al. 1986). The auditory effect of such *rises* in Australian and American English is described by Cruttenden as ‘casual’ (1986: 143).

Turning now to regional variation in German, similar differences, involving *falls* and *rises*, also appear to exist here. While research is scarce, a number of scholars have called attention to the strong differences between the intonation of northern and southern German dialects. According to E. Sievers (1893), for instance, North and South German intonations are like ‘mirror images’ of each other (‘Where the pitch rises in one, it falls in the other, and vice-versa’)⁶⁹, and according to Gibbon (1998: 93), the intonational patterns of north German coastal areas bear a strong similarity to those of north-eastern Britain. To name one example for a non-standard use of *rises* in Southern Germany, the widespread presence of *rising* nuclear tones in statements is characteristic of the Swiss German Dialect of Schwyzerdeutsch. According to the observation of this writer, this gives rise to a feeling of casual friendliness and openness, similar to the auditory effect which Cruttenden associates with this phenomenon in American and Australian English. The contrast between the intonation of Northern and Southern Germany also applies to yes-no questions. Thus, in her studies of intonational isoglosses in the Lorraine and the Palatinate regions, Güntherodt (1969, 1971, 1973) finds confirmation for the geographical division of question intonations - a predominance of *falling* patterns on yes-no questions in southern areas - which was first proposed by Moulton (1962) and Kufner (1962). According to Schädlich & Eras (1969), a similar north-south division with respect to intonation exists in the eastern areas of Germany⁷⁰.

The signaling of speaker personality

The link between intonation and attributions of speaker personality is an area of research which enjoyed a high degree of popularity in early social psychology and Language & Gender studies. Within the first of these two disciplines, the strong link between intonation and judgements of both attitude and personality, as described in early analyses of English and German intonation (Pike 1945, Hadding-Koch 1956 and von Essen 1964), appears to have been firmly established by two studies conducted by E. Uldall in the early 1960s (Uldall 1960, 1964). Among other points of interests, Uldall’s findings indicate, for instance, that the attitude of ‘authoritativeness’ is signalled by a frequent presence of *final falls*, while *rises*, in particular those ending high, are associated with ‘submissiveness’. While Uldall’s results relate most immediately to speaker attitude, further research in social psychology led to the realization that the long-term, or indeed permanent, display of attitudes such as authoritativeness almost automatically incur personality judgements of dominance, while the display of submissiveness lead to personality judgements of timidity and insecurity. (Scherer 1979). This evidences the fundamental association between *falls* and *rises* and the personality traits of dominance and submissiveness⁷¹.

The interest of Language & Gender studies in the link between intonation and personality appears to have been awakened by early observations on female communicative habits, one example being Pike’s (1945) remark that women are ‘primary users of patterns with final (high) rises’ (1945: 59 f.). Among the first studies on male and female uses of intonation are Brend (1972/1975) and Lakoff (1975), both observing independently that women’s speech typically contains a higher proportion of *rising* nuclear tones, especially (high) rises, than men’s speech⁷². According to Lakoff, this frequent female use of (high) rises, gives rise to

⁶⁹ Quoted in Fox 1984: 4.

⁷⁰ A re-evaluation of Güntherodt’s findings has recently been made by Peters (2001).

⁷¹ As Scherer (1979) observes: ‘The semantic components of *certainty/dominance* and *uncertainty/deference*, which in their original linguistic contexts usually refer to certain ephemeral topic-contents, are erroneously interpreted in terms of long-term or even permanent features of the speaker’s language, this in turn leading to their association with this personality’.

⁷² Similar findings have been reported for the speech of (Austrian) German women by Moosmüller (1997).

personality judgements of ‘uncertainty’, ‘lack of self-assertion’ and ‘submissiveness’ and is therefore to blame for the negative way women are viewed in society. As she notes:

‘One [...] consequence is that [this speech behavior] is taken to reflect something real about character and plays a part in [society’s] not taking a woman seriously or trusting her with any real responsibilities, since she [apparently] can’t make up her mind and isn’t sure of herself’ (1975:17).

While Lakoff’s observations were no doubt true at the time and also highly compatible with the findings of social psychology, her interpretations were opposed by the majority of scholars working within Language & Gender. Follow-up research, searching for alternative functions of (high) *rises* not relating to attitude or personality, showed that these may also serve to communicate ‘non-finality’ and ‘incompleteness’, the linguistic functions of these patterns discussed above (McConnell-Ginet 1978, 1983). In the 1980s, interest in women’s communicative habitus shifted to the discursal use of women’s intonation⁷³ and tone of voice, including other related speech phenomena, such as overall pitch, pitch variability and voice quality⁷⁴.

2.2.3 The communicative functions of nuclear pitch-range

Throughout the tonetic literature the consensus appears to be that *pitch-height* and *-range*, signal primarily attitudinal, i.e. paralinguistic, meaning. However, while the attitudinal labels which appear most frequently relate to the different degrees of ‘strength’, ‘intensity’ or ‘forcefulness’ of utterances, the interpersonal labels interspersed in the literature indicate that, like nuclear direction, nuclear pitch-range can also communicate speaker attitude, as for example towards the listener⁷⁵. As a second function, studies have also revealed that the different types of nuclear pitch-pattern created by variations in *pitch-range* may also be used to signal extralinguistic information of the kind discussed above.

Before beginning the discussion of their different functions, a brief presentation of the nuclear pitch-patterns described in the literature for the English and German language, a point omitted hitherto, is in order⁷⁶.

2.2.3.1 Types of nuclear pitch-pattern in English and German

In the following discussion of the English nuclear pitch-patterns, we begin with the different positions taken by Tench (1995) and Pike (1945), after which their two models will be taken together. For the German language, the positions found in Fox (1984) and Pheby (1975) will be discussed briefly.

It was seen above in Section 2.1.3.2 that while most other (tonetic) analysts of intonation, such as Tench (1995), distinguish between three types (*high*, *mid* and *low*) of *falls*, *rises* and *rise-falls*, Pike’s (1945) model is based on four different pitch-heights for *falls* and *rises*, ranging from level 1 to level 4⁷⁷. As Pike allows for greater variation in pitch-range and also does not correlate nuclear *pitch-height* with *pitch-range*, his model contains a larger number of *falls* and *rises*, including *high falls* falling to *mid-high* (i.e. ranging from level 1 to 2) and falling to *mid-low* (i.e. ranging from level 1 to 3). His model also includes *low rises* rising to *high* (i.e. ranging from level 4 to 1) as well as to *mid-high* (i.e. from level 4 to 2) and ‘*mid to high rises*’ (i.e. ranging from level 2 to 1). As another important point, Pike also observes that

⁷³ Perhaps the most important observation to grow out of this discussion was that intonation can serve many different functions. For discussions of this and related themes, see Edelsky 1979, McConnell-Ginet 1978, 1983, Thorne et al. 1983 and Coates & Cameron 1989.

⁷⁴ Some of the research results of Language and Gender studies with respect to these features will be mentioned below, in the sections discussing these features.

⁷⁵ These are often found side-by-side with non-interpersonal attitudinal labels. For example, in Halliday 1970, where the attitudinal meanings of intonation are collectively referred to as the ‘key’ of utterances, this is first defined in terms of ‘the speaker’s attitude of politeness, assertiveness, indifference etc.’ (p. 22), then, a little further down, as ‘the degree of forcefulness or emotional intensity of the utterance’ (p. 31).

⁷⁶ The following discussion is based on Tench 1995 and Pike 1945 for English, and Fox 1984 for German.

⁷⁷ In Pike’s framework, pitch level 1 is the highest, level 4 the lowest level.

falls and *rises* covering a small range of one level generally communicate a feeling of ‘mildness’, pitch-patterns with wider ranges, such as high falls to low (1 - 4) and the low rises to (mid-)high (4/3 - 1) have a ‘strong, intense’ feeling about them. The large variety of *falls* and *rises* in Pike’s framework is illustrated in Fig 10⁷⁸:

High fall to low (1 - 4) completeness/finality
 High fall to mid-low (1 - 3) non-finality
 High fall to mid-high (1 - 2) lightness
 Mid(-high) fall to low (2 - 4) moderateness

Mid-low rise to high (3 - 1) incompleteness
 Low rise to high (4 - 1) deliberative
 Mid(-high) rise to high (2 - 1) politeness
 Low rise to mid-low (4 - 3) incompleteness
 Low rise to mid(-high) (4 - 2) (in sequences)
 Low or mid rise to high (4/3 - 1) intenseness

Fig. 10 Pike’s (1945) *falling* and *rising* nuclear pitch-patterns

However, as comprehensive as Pike’s description is for *falls* and *rises*, on the other hand, his model only allows for a very small number of bidirectional pitch-patterns: one single *fall-rise* and one *rise-fall*. With respect to these pitch-patterns, therefore, Tench’s (1995) model is more precise, differentiating as it does between three (high, mid and low) *fall-rises* and the same number of *rise-falls*⁷⁹. For a precise and comprehensive description - and analysis - of English intonation, therefore, both models are unified here, the resulting framework shown in Fig. 11, together with the attitudinal functions given by the two analysts - to be discussed below.

Regarding the German nuclear pitch-patterns, Fox (1984) differentiates between three (high, mid and low) types of *fall* and *rise-fall*, but only provides for low and mid *rises* and *levels* (1984: 22f.). The number of nuclear pitch-patterns he provides for is thus quite limited. Interestingly, Pheby (1975) differentiates between three (mid, high and low) *falls*, *rises*, *fall-rises* and *levels*, but provides for only one *rise-fall*, starting at low-mid pitch (1975: 66). If one takes both models together, however, one receives approximately the same number of pitch-patterns for German intonation as are described by Tench (1995) for the English language⁸⁰.

2.2.3.2 The paralinguistic functions of nuclear pitch-patterns

In the following discussion, the paralinguistic functions of the English nuclear pitch-patterns will be largely adopted from Tench (1995), with additions from Pike (1945). The functions suggested by these two analysts relate most immediately to what they call the ‘force’ or ‘strength’ of the message, corresponding to what is referred to here as the speaker’s attitude towards his message. However, Tench’s discussion also includes a number of interpersonal labels adopted from other analysts, such as Palmer (1922), O’Connor & Arnold (1973), Halliday (1970) and Gimson (1989). Finally, Pike also provides interpersonal labels for a number of the nuclear pitch-patterns not found in Tench’s model. The discussion of the paralinguistic functions of the German nuclear tones will be based in the main on Fox (1984)⁸¹.

⁷⁸ Fig. 10 is adopted from Tench’s discussion of Pike’s model (1995: 118f.). The conversions from Pike’s 4 levels to the three levels of other tonetic descriptions also follow Tench’s interpretation, according to which, for example, Pike’s mid-high level 2 corresponds to the ‘mid’ level of frameworks with three levels.

⁷⁹ In Tench’s model, all *mid* pitch-patterns are regarded as attitudinally neutral.

⁸⁰ As noted above, Pheby’s (1975) approach to intonational meaning is different from that adopted here, so that his meanings only receive limited treatment here.

⁸¹ In those cases where no interpersonal labels are to be found in the literature, these meanings will be derived from those discussed in other studies. These are either non-interpersonal labels given for the German pitch-patterns or the interpersonal labels of their English counterparts.

*Tench's attitudinal meanings (1995)*⁸²

In Tench's model all *mid* nuclear tones are regarded as neutral, signaling a neutral attitude on the part of the speaker. The *high fall*, falling from high to low, is associated with 'intensity' and 'strength', while the *low-fall*, ranging from mid-low to low pitch, is given the labels 'mild', 'detached' and 'reserved', the latter two clearly relating to interpersonal attitude. With regards to the interpersonal meanings of the *high fall*, 'insistence' and 'dominance' seem appropriate and will be adopted here. Coming to the *rises*, Tench's *high rise*, ranging from low to high, is given the meanings of 'intensity', 'concern' and 'eagerness', the latter two more interpersonally oriented than the first. With respect to the *low rise*, Tench observes that it is often associated with rather negative labels, ranging from 'reserved', and 'resentful' to the more neutral term 'deliberative', which he ultimately adopts⁸³. His interpersonally oriented labels 'reserved', and 'possibly negative' are also adopted here. Regarding *fall-rises*, Tench distinguishes between *low* and *mid* variants, associating the first with a high, the second with a lower degree of 'contrastiveness' and 'implication'. The interpersonal meaning of 'assertiveness' is thus proposed here for this pitch-pattern. Finally, Tench associates the *high rise-fall* with 'intensity' and 'assertiveness', while the *low rise-fall* is described as signaling 'intensity with a heightened degree of emotional involvement'. The label 'intensity plus emotion' is also adopted from Tench for this pitch-pattern.

*Pike's attitudinal meanings (1945)*⁸⁴

Three of Pike's pitch-patterns, not included in Tench (1995), are of special interest to us here: the *mid-fall*, the *high narrow fall* and the *mid-rise*. While the *mid-fall*, ranging from mid/mid-high to low, is associated with a 'gentle' and 'moderate' feeling, the most gentle and pleasant labels are reserved for the *high narrow fall*, extending from high to mid-high. This variant is described as signaling a 'mild', 'light', 'endearing' and 'polite' feeling, the latter two labels relating to interpersonal attitude. Pike also observes that the *high narrow fall* is often used in 'baby talk', a manner of speaking adopted by adults when speaking to children⁸⁵. Tench's *mid-high rise*, rising from mid to high pitch, is also associated with the feeling of 'mildness', 'gentleness' and 'politeness', and Pike notes that if used excessively, this pitch-pattern may be interpreted in terms of 'unctuousness' and 'insincere politeness'⁸⁶. In the present description, Pike's labels for these three variants are maintained, but the term 'endearing' is replaced by 'warm'. The above described attitudinal meanings of the English nuclear pitch-patterns are summarized below in Fig. 11.

Coming now to the German nuclear pitch-patterns, not much is to be found in the tonetic literature in the way of interpersonal meanings. The functions suggested by Fox (1984: 58 ff. and 67 f.) are of a very broad nature, relating simply to the degree of 'involvement' or 'intensity' signaled by the utterances. Similar to what was found above for English intonation, nuclear pitch-patterns with a wider range are widely associated with 'strength' and 'intensity', while narrower pitch-patterns are associated with 'weakness' or 'unemphaticness'. Unfortunately, Fox does not discuss the attitudinal functions of particular types of nuclear pitch-patterns, so that the validity of attitudinal meanings discussed above for the English nuclear pitch-patterns with respect to their German correlates is unclear. On the other hand, the attitudinal meanings associated with pitch-range are clearly the same, so that it seems possible to apply the attitudinal and interpersonal meanings discussed above for the English nuclear pitch-patterns to the German patterns discussed by Fox and Pheby⁸⁷.

⁸² The following discussion is based on Tench 1995: 115 ff.

⁸³ Cruttenden suggests the similarly negative connotations of 'mild unpleasantness' and 'grouchiness' (1986: 106).

⁸⁴ The discussion of Pike's meanings (1945: 44 ff.) is based on Tench's interpretations (1995: 115 ff.).

⁸⁵ Crystal 1987: 416. Another term for 'baby talk' is motherese. (See Chapter 5)

⁸⁶ The attitude Pike is referring to here could conceivably correspond to that communicated in *false deference*. (See Chapter 1)

⁸⁷ Since much the same meanings appear to be associated with pitch-height and accent-range in English as discussed above for German, the applicability of these meanings seems strong. (See also the further results of Uldall's studies discussed below)

<i>Nuclear pitch-pattern</i>	<i>Attitudinal functions</i>	
	<i>towards message</i>	<i>towards listener</i>
High fall	intensity, strength	insistence, dominance
Low fall	mild	detached, reserved
Mid fall	gentle, moderate	gentle
High fall to mid-high	mild, light	warm, polite
High rise	intensity	concern, eagerness
Low rise	deliberative	reserved, poss. negative
Mid (-high) rise	mild	gentle, polite
High fall-rise	strong contrastiveness	strong assertiveness
Low fall-rise	moderate contrastiveness	assertiveness
High rise-fall	strong intensity	strong assertiveness
Low rise-fall	intensity	intensity + emotion

Fig. 11 The attitudinal functions of the English nuclear pitch-patterns

2.2.3.3 The extralinguistic functions of nuclear pitch-patterns

Though research in this field is rare and also largely limited to the English language, there is evidence that extralinguistic information - relating to speaker personality and speech community membership - can also be communicated by nuclear pitch-range, both in English and in German. With respect to speaker attitude/personality, for example, E. Uldall's (1960, 1964) study of English intonation mentioned above also found that *mid rises* are associated with 'authoritativeness', while *high rises* clearly communicate 'submissiveness'. Her results also suggest a strong link between wide accent-range and 'authoritativeness', between narrow accent-range and 'submissiveness' (1960: 233)⁸⁸, a finding in agreement with Pike's (1945) observations on this feature.

As a second area of research, studies in comparative English intonation have also examined the extralinguistic functions of certain types of nuclear pitch-pattern. While the communicated information relates most immediately to speech community membership, the frequent use of certain nuclear pitch-patterns often typical of certain socio-regional variants of English also give rise to attitude and personality judgements on the part of speakers of other varieties. To name one frequently cited example, Americans have been found to use *high rise-falls* more often than speakers of British English and also appear to prefer *high-rises* in situations where *low-rises* would be expected in R.P. (Pike 1945, Cruttenden 1986). Regarding the effect of such differences in communicative habitus, Bolinger (1989), commenting on the British English preference for these 'sudden' *falls*, notes that to American ears, British speaker give the impression of 'not merely reacting to the situation but controlling it' (1989: 31).

Similar findings have been reported in studies of intercultural communication. Gumperz, for example, notes that speakers of Indian English tend to use *high falls* and *high rise-falls* in situations where speakers of (western) British English typically employ attitudinally neutral *mid falls* (Gumperz 1982 a, b).⁸⁹ According to his observations, a predominant use of these pitch-patterns creates a strong impression of 'rudeness' and 'pushiness' in native speakers of British English, a finding also reported by Bolinger (1989: 62 f.)⁹⁰. As another example from this area of research, Kvavik (1988) reports that Spanish speakers of English frequently use *low falls* instead of *mid* or *high falls* when answering questions. In Bolinger's interpretation, this is likely to sound 'rudely assured' to speakers of R.P. (1989: 63).

Though research in this area regarding the German language is rare, it appears that here too, regional preferences for certain types of nuclear pitch-pattern not only exist but are also frequently associated with much the same - usually negative - attributes of attitude and

⁸⁸ As can be seen, Uldall's findings are highly compatible with the attitudinal meanings suggested by Pike (1945) for mid and high rises and with Pike's and Fox's association between narrow accent-ranges with 'gentleness', wide accent-ranges with 'strength'.

⁸⁹ Similar reports are found in Bailey 1983, and Bolinger 1989.

⁹⁰ See also Gumperz 1982a: 168.

personality as their English counterparts. To name one frequently cited example, it appears that north German dialects are characterized by wider and also more steeply descending *falls* than southern dialects, a feature frequently misinterpreted in terms of arrogance and superiority by speakers of south German dialects⁹¹. According to the observation of this writer, wide and abruptly descending *falls*, giving rise to precisely this auditory impression with speakers of other regional variants of High German, are also typical of the variant spoken in the Saarland area. As will be seen in Part Three (Chapter 9), such wide steep *falls* are relatively frequent in the German dialogs recorded in the experimental part of this work⁹². Other regional idiosyncracies of German intonation are discussed in Gibbon 1998.

2.2.4 The communicative functions of tail-slope configurations

As will be recalled from Section 2.1.2.2, Couper-Kuhlen (1986) distinguishes between three English tail-slope configurations which she associates with different attitudinal (non-interpersonal) meanings⁹³. Thus, according to her analysis, gradual, gentle, tail-slopes represent the norm, so that no particular attitudinal message is associated with this variant. As the second type, steep tail-slopes, often found in wide nuclear pitch-patterns, are described as signaling ‘contrastivity’, associable with strength and forcefulness. Finally, *flat tails*, in which the pitch-heights of all the tail syllables have approximately the same, usually low pitch-height, are given the labels ‘boredom’, ‘sarcasm’ and ‘monotony’. In this context Couper-Kuhlen adds that *flat tails* are also a feature of *stylization*, a particular type of contour consisting of two or more consecutive level pitch-patterns - generally a low level *head* followed by a mid or high level nuclear tone and a *flat tail*. According to Ladd (1980) who first suggested the term of ‘*stylized fall*’, this pattern signals ‘routineness’, ‘predictability’ and ‘stereotype’⁹⁴.

Tail-slope configuration	Attitudinal functions	
	towards the message	interpersonal attitude
Gentle tail	moderateness	gentleness
Steep tail	contrastivity	strong position, dominance
Flat tail	monotony	boredom, sarcasm

Fig.12 The attitudinal functions of English tail-slopes

The meanings Couper-Kuhlen suggests for steep tails (contrastivity) and *flat tails* (‘boredom’, ‘sarcasm’ and ‘monotony’) are maintained here, but for steep tails the more interpersonally oriented labels of ‘strength’ and ‘dominance’ are added to the meanings given above. Gentle tails will receive the labels ‘moderateness’ and ‘gentleness’. These meanings are summarized in Fig. 12 above.

2.2.5 The functions of the prenuclear pitch-patterns

The prenuclear pitch-patterns are also associated with attitudinal meanings in the literature. With the exception of the *stepping* and *glissando* types, however, these meanings relate less to interpersonal attitude than the manner in which the speaker presents his message. As a second function, the prenuclear pitch-patterns are also ascribed extralinguistic functions.

⁹¹ Klaus Kohler, personal communication. See also Kohler 1995.

⁹² See Chapter 9, where these pitch-patterns will be referred to as *quasi-stylization*.

⁹³ As Couper-Kuhlen discusses only English tail-slopes and this feature is not included in tonetic descriptions of German intonation, the functions of different tail-slopes in German, though assumed to also exist in this language, is not discussed here. This is a further issue to be examined in Part Three (Chapter 9).

⁹⁴ Ladd 1980: 266, quoted by Couper-Kuhlen 1986: 88. Note that these meanings are highly reminiscent of those discussed above for non-final level tones in German: ‘triviality’ and ‘self-explanatoriness’ and the interpersonal meanings of ‘disinterest’ and ‘boredness’.

2.2.5.1 The paralinguistic functions of the prenuclear pitch-patterns⁹⁵

As discussed in Section 2.1.3.1, Tench distinguishes between three types of prenuclear pitch-pattern: level *heads*, including gently falling and rising patterns and wide ascending or descending *heads*. The level *heads* are divided into three types, *high*, *mid* and *low*, the mid variant being regarded as attitudinally neutral⁹⁶. The meanings of the other two level *heads* depend to a certain degree on the type of the nuclear pitch-pattern following in each case: Thus, low level *heads* before high and low rising nuclear pitch-patterns are seen to signal a high degree of ‘involvement’, the first pattern additionally associated with a ‘warm’ feeling, while low level *heads* before *low rises* signal ‘coldness’ with a reproachful and somewhat menacing undertone. High level *heads* before *high falls* receive the labels ‘insistence’ and ‘involvement’, and high level *heads* before *low falls* are associated with ‘weightiness’, in the sense of ‘emphaticness’. Finally, low level *heads* before falling nuclear pitch-patterns and high level *heads* before *rises* are seen as ‘focusing on’ or ‘highlighting’ the content of the speaker’s message.

Type of head	Type of nuclear pitch-pattern	Attitudinal meaning
High level	rises	focusing, contrastive pointing
Low level	falls	focusing, contrastive pointing
Low level	high rise	involvement + warmth
Low level	low rise	involvement + coldness
High level	high fall	involvement, insistence
High level	low fall	emphaticness, weightiness
wide descending	(not specified)	authority + warmth
wide ascending	(not specified)	appeal + warmth
stepping head	(not specified)	emphasis
glissando head	(not specified)	insistence, forcefulness

Fig. 13 The attitudinal functions of the English prenuclear pitch-patterns (Tench 1995)

The meanings of Tench’s wide descending/ascending *heads* depend on whether they are simple falls/rises or contain *stepping* or *glissando* movements. Generally speaking, the wide ascending *head* is seen to signal an ‘appeal to the listener’ combined with a ‘warm’ feeling, while the wide descending *head*, though also carrying a warm undertone, is given the labels ‘authority’ and ‘insistence’. Turning now to the *stepping* and *glissando* types, these are seen to add increasingly strong degrees of ‘strength’ or ‘insistence’ to the basic meanings of ascending and descending *heads*. Thus, compared with the more neutral matter-of-factness of simple *rises* and *falls*, *stepping heads* are associated with ‘emphasis’, while the *glissando* type, the strongest of all, receives the labels ‘insistent’, ‘forceful’, and ‘conceding no opportunity for contradiction [by the listener]’ (1995:133). Tench’s meanings of the English prenuclear pitch-patterns are summarized in Fig. 13.

Coming to the German prenuclear pitch-patterns, Pheby (1975) and Fox (1984) associate their less numerous *heads* with different degrees of emphasis and/or forcefulness, these meanings again relating less to interpersonal attitude and more to the way a message is presented. Like Tench, Pheby links the meanings of his prenuclear pitch-patterns to the nuclear patterns following in each case and also postulates attitudinally neutral forms. Thus, level and simple falling *heads* preceding wide falls are regarded as neutral, while falling *heads* in this position signal ‘simple emphasis’ and rising *heads* before *falls* are associated with ‘expressive emphasis’. With rising nuclear patterns, narrow (simple) falling *heads* are regarded as neutral, wide (simple) *falls* associated with a ‘sceptical’ undertone. In Fox’s discussion the meanings

⁹⁵ The following discussion is based on Tench 1995:128f., Pheby 1975: 59f. and Fox 1984: 23f.

⁹⁶ In fact, since Tench regards all mid nuclear pitch-patterns as attitudinally neutral, not only the mid level *heads* but also all *heads* preceding mid nuclear pitch-patterns are associated with ‘neutralness’ and not discussed in more detail. Following Pike’s treatment of mid falls and mid rises, the labels of ‘mildness’ and ‘moderateness’ are associated with these patterns here.

are assigned in a more general manner: Low and high *heads* are seen as conveying the ‘relative importance’ of the speaker’s message. While high level *heads* with falls communicate the ‘importance’ of the speaker’s message, low *heads* with falls are regarded as a kind of neutral variant, implying that the message may be expected to contain only ‘known’ or ‘expected’ information. Low *heads* preceding rises are described as having ‘contrasting’ and ‘pointing out’ effect. Finally, the rising *head* is associated with ‘strong emphasis’.

2.2.5.2 The extralinguistic functions of the prenuclear pitch-patterns

Like the nuclear directions and the different types of nuclear pitch-pattern discussed above, English⁹⁷ prenuclear pitch-patterns can also signal speaker affiliation with certain socio-regional varieties. In addition, studies have shown that the prenuclear pitch-patterns typical of some regional variants of English are also frequently associated with certain - once again largely negative - personality traits by speakers of other dialects. The strong variation found in different English varieties with respect to prenuclear patterns is illustrated in Fig. 14, adopted from Bailey 1983, which shows the prenuclear pitch-patterns described as typical of R.P., Irish English, and Northern/Southern General American. According to Bailey’s analysis, members of other speech communities are likely to find the English R.P. version ‘condescending’, the Irish ‘insultingly incredulous’, the Northern American ‘repetitive’, and the Southern American ‘unaccountably surprised’ (Bailey 1983: 17).

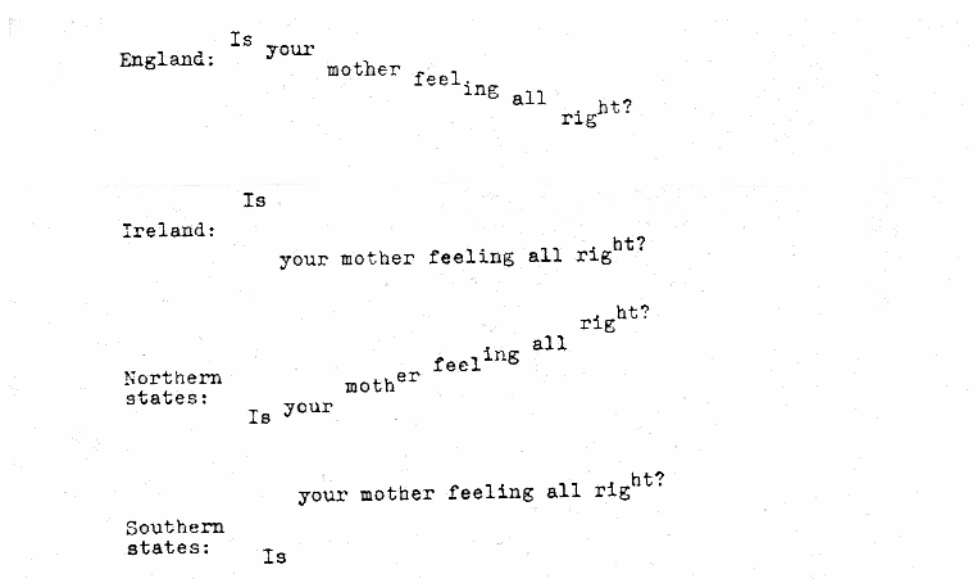


Fig. 14 Four prenuclear pitch-patterns of English (after Bailey 1983)

⁹⁷ Unfortunately, no data were available to this writer on this topic with regards to the German language.

3. The communicative forms of Mandarin Chinese intonation

Chinese intonation research is still at an early stage of development. Despite a recent surge in interest in Chinese suprasegmental communication, felt most strongly in the field of speech technology, progress in the study of Chinese intonation remains slow. Researchers are still grappling with fundamental problems, such as the workings and interplay of the lexical tones (*tonation*), the influence of emotion on the phonetic *gestalt* of lexical tones and the - phonetic or phonological - status of pitch-declination in Chinese. As a result, only very little is known about the manner in which intonation is manifested in Mandarin Chinese, i.e. its communicative *forms*, and even less of their communicative functions, so that a detailed discussion of the *forms and functions* of Chinese intonation of the kind undertaken above for English and German is not yet possible at this point. As a first step towards such an understanding, an attempt will be made here to define the communicative forms of Mandarin Chinese intonation on the basis of the existing research literature, notably Chao 1968, Shen 1990, Tao 1996 and Kratochvil 1998, and the evidence found in the speech corpus compiled for this work. On the basis of this evidence, the communicative functions suggested for these forms in the literature will also be discussed¹. Before beginning with the discussion of Mandarin Chinese intonation, a brief presentation of the Mandarin lexical tones is in order, focusing on their basic configurations and the way these are modified by certain phonetic and phonological processes.

3.1 The Mandarin Chinese tones

3.1.1 The lexical tones in isolation - basic configurations

Like many of the languages spoken in Southeast Asia, Mandarin Chinese is a tone language, with 'lexically significant, contrastive', and thus phonological, pitch-patterns on each syllable (Pike 1948:3)². In the languages and dialects of China the phonological use of pitch is manifested in the *lexical tones*, the characteristic pitch-patterns serving to distinguish different words/morphemes³. These lexical tones, or *tones*, are as much part of the phonological make-up of Chinese morphemes as consonants and vowels. Tones may be level or contoured, the latter involving falling or rising movements, or a combination of the two.

In contrast to the large tonal sets of other Chinese languages, such as Cantonese and Min, Mandarin Chinese has only four lexical tones, numbered as the 1st, 2nd, 3rd and 4th tones. Their configurations are described in terms of two criteria: (1) the shape of the pitch-movement, i.e. *level*, *rising*, *falling* or combinations of the latter two, and (2) their (initial) pitch, in terms of the position of their starting points within the speaker's voice-range, divided into the five levels *high*, *mid-high*, *mid*, *mid-low* and *low*. Using these criteria Tone 1 is described as *high level*, Tone 2 as *mid* or *mid-high rising*, Tone 3 as *low fall-rising*, or *low dipping*, and Tone 4 is defined as *high falling*. Fig 11, adapted from Chao (1968), shows the basic configurations of the four Mandarin tones relative to each other, represented as notes on a musical stave. The five voice-range levels are also represented with small numbers. Chao's musical stave is referred to here as the *tonal-band*⁴, and the breadth of this musical stave, representing the pitch-range between the 1st and 5th level at any given point in an utterance, is referred to as *tonal band-width*. The 1st and 5th levels are also referred to as the *bottom* and *top line* of the *tonal band-width*. While there is no agreement on the - absolute or relative -

¹ An in-depth examination of the communicative functions of the Chinese intonation forms suggested here, in particular those relating to interpersonal attitude, is undertaken in Part Three. See Chapters 9 and 10.

² For a detailed discussion of *tone* and tone-languages, see the papers collected in Fromkin 1978.

³ Mandarin Chinese is basically monosyllabic, so that each syllable represents one morpheme. (Li & Thompson 1989, Norman 1988, Chao 1968). However, in addition to the relatively small, closed set of monosyllabic morphemes, most of which may be traced back to ancient pre-Han times, the modern standard language also has a large number of di-, tri- and even quadrisyllabic morphemes, formed by combinations of single morphemes. These multisyllabic words are steadily increasing as more and more new words are created to cope with the growing number of new concepts of modern life.

⁴ The writer is indebted to Paul Thompson of the London School of Oriental and African Studies for having suggested this term.

dimensions of the intervals which separate the five *levels*, it is known that they vary with degree of emotional involvement (Chao 1968: 26)⁵.

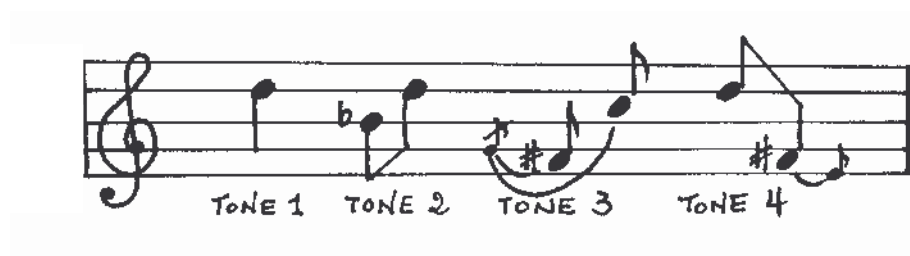


Fig. 15 The configurations of the four Mandarin tones (Chao 1968)

As observed by Kratochvil (1998), not only pitch-movement and pitch-height but also relative amplitude - in terms of amplitude variation over the course of a tone - serve as distinctive features of the lexical tones⁶. Fig.16, based on the research results discussed in Kratochvil (1998: 421f.), shows the configuration of the four Mandarin tones as a function of relative height, pitch-movement and relative amplitude⁷. As Fig.16 illustrates, the tones differ quite considerably with respect to the position of their points of highest intensity (represented here as large black dots) over their length: While in Tones 1 and 4 these points roughly coincide with their starting-points - allowing for a brief on-glide - this is not the case in Tones 2 and 3, whose points of highest intensity come well after the beginning of their pitch-movements⁸. Based on Fig. 15 and 16, it seems possible to say that the loudest points of Tones 1 and 4, corresponding roughly to their initial pitch/frequency, coincide with the *top line* of the tonal band. In contrast, the point of highest intensity of Tone 2 occurs between the *mid* and *mid-high levels* of the tonal band, while the point of highest intensity of Tone 3, coming after it makes its upward turn, coincides with the *bottom line* of the tonal band-width. Regarding the final pitch- or frequency points of the four Mandarin tones, though these vary as a function of the pitch-range covered by the tones, it seems possible to say that the ending-point of Tone 2 ends a little below the *top line* of the tonal band, while the lowest pitch-point achieved by a 4th tone is roughly associable with the *bottom line*⁹.

The contrastive functions of the four Mandarin lexical tones are illustrated by the following two groups of morphemes¹⁰:

<i>ma</i> ¹	mother	<i>ma</i> ²	hemp	<i>ma</i> ³	horse	<i>ma</i> ⁴	scold
<i>hai</i> ¹	sigh	<i>hai</i> ²	child	<i>hai</i> ³	ocean	<i>hai</i> ⁴	harm

In addition to the four tones discussed up to this point, Mandarin Chinese has a 5th tone, referred to in the literature as the *light* or *neutral tone*. Found only on non-stressed, *neutral* or *neutralized*, syllables¹¹, its configuration is very different: Lacking both a characteristic height and a fixed pitch-movement, it is realized as a brief tap-like sound whose pitch-height is determined by its tonal environment and factors of stress and intonation¹².

⁵ See Chao 1968: 26. More will be said of this phenomenon in Chapter 9. Chao's proposal of a correlation between tonal band-width and what he refers to as the speaker's mood has received strong support from recent phonetic studies, such as Yang & Campbell 2001.

⁶ Kratochvil 1998: 419. See also Coster & Kratochvil 1984.

⁷ Kratochvil's data represent the mean frequency and amplitude values for the four Mandarin tones calculated on the basis of a speech corpus, a monolog of spontaneous speech containing 1390 syllables and produced by a female native speaker of Mandarin (1998: 418).

⁸ To be quite accurate, allowance must be made for what seem to be *on-glides* - falling in Tones 1 and 4, and fall-rising in Tones 2 and 3.

⁹ Strictly speaking, the last two statements apply only to full, non-reduced, 2nd and 4th tones, a point of which more will be said below.

¹⁰ The symbols and diacritics used in the representation of Chinese tones are presented at the end of the General Introduction.

¹¹ These are also referred to as *atonal* syllables. The neutral tone is a characteristic feature of many Chinese particles, suffixes and interjections (See Chao 1968: 32 ff. and Shen 1990: 32, 38 ff.). Apart from this, neutral, or *neutralized*, tones also appear on syllables whose *original*, lexical tones have been neutralized under the effect of prominence, a process to be discussed below.

¹² This issue is discussed in some detail in Shen 1990: 38 ff.

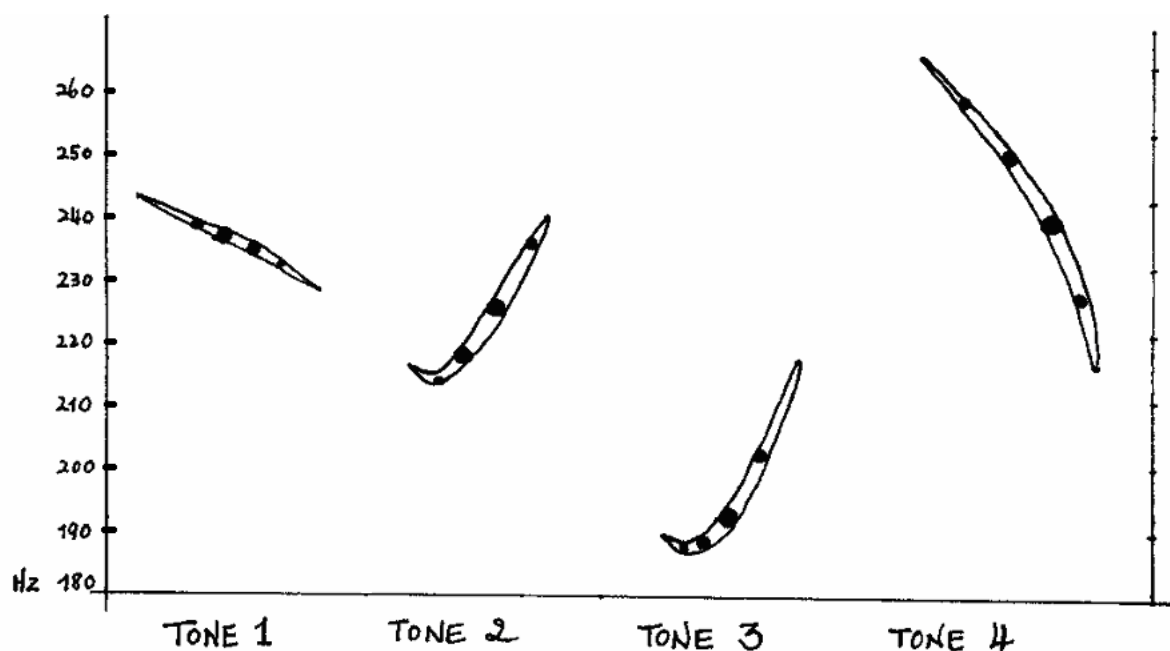


Fig. 16 The configuration of the four Mandarin tones (after Kratochvil 1998)

3.1.2 The tones in tonal compounds

The above description of the four Mandarin tones applies only to their citation forms, spoken in isolation. As soon as syllables combine to form di-, tri- or polysyllabic tonal compounds, representing compound morphemes, such as *lao*^{3/2}-*ban*³ ('boss'), *zai*⁴-*jian*⁴ ('good-bye') and *mai*³-*cai*⁴ ('buy vegetables')¹³, the basic pitch-configurations of the tones are modified according to two sets of phonological rules, the rules of *prominence* and the rules of *tone sandhi*.

3.1.2.1 The rules of prominence¹⁴

In tonal compounds of two or more syllables, or morphemes, prominent and non-prominent syllables alternate regularly¹⁵. Under the effect of prominence Chinese syllables undergo modifications relating to loudness and duration as well as the pitch-height and -range of their tones. Kratochvil describes the effect of prominence as follows:

Phonetically, [prominence] is a relative enhancement of the primary tonal properties [...]. Prominent syllables are generally louder, their pitch level is higher (in the case of tones characterized by a *fo* peak, that is Tones 1, 2 and 4, as well as most atonic syllables) or lower (in the case of Tone 3, characterized by a *fo* trough), and their duration greater, especially in the case of Tone 3, in relation to their non-prominent correlates. [As a result of these modifications], the contours of pitch and loudness curves of prominent syllables appear as exaggerations: the peaks are higher, the troughs are lower, and the slopes towards and away from them are more pronounced. (1998: 422)

In contrast, non-prominent syllables of tonal compounds are marked by a decrease in duration and amplitude and a reduction of the length and range of their tonal pitch-movements: 1st tones become shorter, 2nd and 3rd tones rise less high, and 4th tones fall less deeply than their prominent correlates. As Kratochvil adds, depending on factors such as tempo and style, the degree of modification undergone by the tone of a destressed syllable may vary considerably, ranging from a simple *reduction* of its pitch-movement and range to what is referred to here as *neutralization*, where the pitch-movement of the tone is reduced to zero and the tone adopts the tonal features of the *neutral tone* (1998: 422).

¹³ *Mai*³-*cai*⁴ may also be broadly glossed as 'going shopping'. The phenomenon of the tonal compound is seen here as including (noun- or verb-) phrases, i.e. sequences of morphemes bound by syntactic rules.

¹⁴ As in Chapter 2, the term 'prominence' is preferred here to the term 'stress' commonly used in the sinological literature.

¹⁵ According to Kratochvil (1998), the alternation of prominent and non-prominent syllables in Mandarin manifests a iambic tendency (1998: 423 f. See also Chao 1968: 148 ff.). This does not apply strictly to Taiwanese Mandarin, where the opposite is often the case.

3.1.2.2 The rules of tone sandhi

In most, perhaps all, Chinese languages and dialects, juxtaposition of the lexical tones in connected speech gives rise to a second set of modifications referred to as *tone sandhi*¹⁶. Numerous and highly complex in some Chinese dialects, such as the Southern Min dialect spoken in Fujian and Taiwan, the *tone sandhi* phenomena of Mandarin are few and apply only under certain limited circumstances, as when tonal compounds are formed. In contrast to the comparatively minor changes in pitch incurred by prominence, the modifications which are referred to as *tone sandhi* are of such a scale that affected tones are literally changed beyond recognition¹⁷, their resulting pitch-configurations often coinciding with those of other tones. In Mandarin, this may be illustrated by the *tone sandhi* affecting Tone 3, which applies when it is followed by a further 3rd tone and effectively transforms it into a 2nd Tone¹⁸. This is illustrated in the following four examples involving sequences of two and three syllables:

- (1) $hao^3 + jiu^3 \Rightarrow hao^2-jiu^3$ (2) $mai^3 + hao^3 + jiu^3 \Rightarrow mai^2 hao^2-jiu^3$
 good wine good wine buy good wine buy good wine
- (3) $mai^3 + hao^3 + jiu^3 \Rightarrow mai^3 hao^2-jiu^3$ (4) $wo^3 + hen^3 + hao^3 \Rightarrow wo^3 hen^2-hao^3$
 buy very long time take long time to buy I very good I ('m) very well

As the above examples illustrate, the rules governing the *tone sandhi* changes of Tone 3 are far from simple. While the basic rule is that only syllables belonging to one compound are affected, a fact illustrated in example (1), it seems that in the formation and joining of compounds in coherent speech, the application of *sandhi* is dependent on a number of different factors regulating what might be referred to as the 'affinity', or 'cohesion', between individual morphemes and tones. Thus, in example (2), the fact that the first two 3rd tones - on mai^3 and hao^3 - are both changed to 2nd tones, may be an indication that the three morphemes in the compound $mai^2hao^2-jiu^3$ are bound by the same, strong, degree of cohesion. By contrast, in example (3) the fact that mai^3 remains unaffected by *sandhi* which, however, then affects hao^3 , suggests that the cohesion between mai^3 and hao^3 is looser than between hao^3 and jiu^3 . A number of phonetic and syntactic factors are discussed in the literature. In this particular case, it could also be stress, in terms of *emphasis*, which is involved¹⁹. Another example for the non-application of *sandhi* is shown in (4), where the cohesion between wo^3 and hen^3 is obviously less strong than between hen^3 and hao^3 .

A second *tone sandhi* encountered in Mandarin is that of Tone 2. However, as it is limited to trisyllabic tonal compounds, it is generally regarded to be of minor importance. Following Chao (1968:27f), the rule is formulated as follows: 'If in a three-syllabic word or phrase ABC, A is in the 1st or 2nd Tone, B in the 2nd Tone, and C in any except the neutral tone, then B changes into the 1st Tone for speech at conversational speed, but does not change at a more deliberate speed'. Among the examples Chao gives for this modification are the following:

- (1) $san^1 + nian^2 + ji^2 \Rightarrow san^1-nian^{2/1}-ji^2$ (2) $hai^2 + mei^2 + wan^2 \Rightarrow hai^2 mei^{2/1} wan^2$
 three year grade third grade yet not finished not yet finished

While other *tone sandhi* changes, relating to the 3rd and the 4th Tones, are sometimes discussed in the literature, a close look reveals that they are in fact explainable as modifications

¹⁶ The term *sandhi* is adopted from Indo-European studies, where it is used to denote certain assimilatory processes affecting morpho-phonological forms in juxtaposition. (See Robins 1980, Szemerényi 1989)

¹⁷ In view of this difference, Norman (1988: 146) suggests a distinction between 'phonetic' and 'phonemic' *sandhi*.

¹⁸ In the literature this is the most commonly encountered formulation of the rule. Note, however, that in Kratochvil's opinion, the tonal product of 3rd tone *sandhi* is not a 2nd tone, but 'an idiosyncratic form distinct from both Tone 3 and Tone 2' (1998: 425). This view is also expounded in Kratochvil (1987) and Zee (1980).

¹⁹ Discussions of phonetic factors involved in these processes are found in Shen 1990. The influence of stress does not seem to have received much consideration yet.

due to prominence. A case in point is the ‘tone *sandhi* rule’ affecting the 4th Tone, and described by Shen (1990:33) as follows: ‘When a 4th tone is followed by another 4th tone, the first does not fall to the bottom, but only to the middle’. The example she gives is *zai*⁴ - *jian*⁴ (‘good-bye’). A closer look, however, reveals that this process corresponds exactly to what was described above as the reduction of the pitch-movement and pitch-range of a 4th tone due to destressing within a tonal compound²⁰. As a last point, it should perhaps be added that tone *sandhi* changes apply before stress-induced modifications²¹.

We now come to the discussion of Chinese sentence intonation, which is divided into three parts. As in Chapter 2, we begin with Chinese *tonality* and *tonicity*, after which we will take a look at intonational pitch-patterns in Mandarin Chinese.

3.2 Tonicity in Mandarin Chinese

This section begins with a brief discussion of the manner in which *tonality* is manifested in Mandarin Chinese, in terms of the subdivision of Chinese utterances into units of tonality. Following this, we take a look at how the boundaries of these tonality units are signaled by length, loudness, pitch and pause.

3.2.1 The Mandarin Chinese intonation-group

As in western languages, the highest functional unit of Chinese intonation is the intonation-group (Chao 1968, Shen 1990, Tao 1996). Similar to what was seen above for English and German, Chinese intonation-groups often coincide with sentences and clauses, but can also be found to coincide with clause constituents²². Greenbaum & Quirk’s (1990) distinction between major and minor intonation-groups can also be applied to Mandarin Chinese: As in English and German, major intonation-groups are defined here as coinciding with whole sentences, while minor intonation-groups, separated by minor intonation-group boundaries, are the intonational equivalents of clauses²³. On the lowest, more local level, the basic unit of Chinese intonation is agreed to be what is called the *prosodic word*²⁴ in the phonetic literature. To be referred to here as the *tone-unit*, this unit corresponds to the pitch-pattern of one morphemic unit - either a single morpheme, bearing a single tone, or a compound morpheme, in which two or more consecutive tones form a *tonal compound* in the manner discussed above. As such, the *tone-unit* contains by definition at least one but normally not more than one prominent syllable²⁵. On the intermediate level a third unit is proposed here, the *tone-group*. Coinciding with clause constituents, such as adverbial clauses, subjects and topics, the *tone-group* consists of one or more *tone-units* bound to each other by syntactic processes²⁶. In the present discussion, therefore, three units of Mandarin Chinese *tonality* are suggested: the intonation-group, the *tone-group* and the *tone-unit*. This is illustrated in the following examples, taken from the Chinese speech corpus recorded for this work²⁷.

- (1) { [[e²-dan⁴] [zhu³-tang¹]] [[bu²] [tai⁴ hao³]] ba } ?
 goose egg make soup not very good part.
 Making soup with goose eggs is not such a good idea, is it?

²⁰ For this reason, Chao (1968: 28f) and Norman (1988:146f.) do not include this phenomenon in their discussions of tone *sandhi*.

²¹ The fact that such stress-induced changes - involving the reduction and sometimes even the neutralization of tones, in terms of their tonal pitch-range - occur in every tonal compound - might explain Kratochvil’s above noted findings with respect to *sandhi*-modified 3rd tones.

²² A detailed discussion of the syntactic units which may coincide with intonation-groups in Mandarin is found in Tao 1996: 55 ff. Interestingly, Tao’s results suggest that the percentage of clausal intonation-groups lies around 50% in spontaneous Mandarin speech, a percentage identical to that found for English by Quirk et al. 1964 and Tench 1990. (See also Tench 1995: 32).

²³ Much the same distinction, with a different nomenclature, is made in recent studies in experimental phonetics, e.g. Yang & Wang 2002.

²⁴ This term, originally created for the description of French intonation (See Wunderli 1989, Di Cristo 1998), is commonly used in studies of Chinese experimental phonetics, e.g. Yang & Wang 2002, Yao & Pan 2002.

²⁵ Note, however, that in polysyllabic morphemic compounds, such as *gong*¹-*gong*⁴ *qí*⁴-*che*¹ (‘bus’), and *guo*²-*ji*⁴ *qíng*¹-*nian*²*huo*²-*dong*⁴ *zhong*¹-*xin*¹ (‘international youth activity center’), there is usually a secondary stress in addition to the more marked primary stress.

²⁶ What is referred to here as the *tone-unit* corresponds to Tao’s (1996) non-clausal intonation-groups.

²⁷ These utterances correspond to the speech samples *wDan 4*, *yLu 2*, *wMan 4* (slightly modified) and *yDang 1* of the Chinese speech corpus. The translations given here are slightly less free than those given in Appendix I, Section II. Intonation-group boundaries are represented with curly braces, tone-unit boundaries with square brackets and tone-groups are separated by vertical lines.

- (2) {[[Ran²-hou⁴]] [[zhe⁴-ge⁰]] [[lu³-rou⁴]] [[duo¹ shao^{3/2}]] [[qian²]]}?
 Afterwards this stewed meat how much money
 Well then, how much [was] the stewed meat ?
- (3) {[[Ni³]] [shuo¹ de⁰]] [[shi⁴]] [[zuo⁴ hao³ de⁰]] [[man²-tou²]] | ma⁰ }?
 you talk part. is/are made ready part. dumpling(s) part.
 Are you talking about ready-made dumplings?
- (4) {[[Xin¹-xian¹]]}. {[[Xin¹-xian¹]]}. {[[Dang¹-ran²]] | [[xin¹-xian¹]]}. {[[Xian¹-zai⁴]] | [[dang¹-ji⁴]]}.
 Fresh. Fresh. Of course fresh. Now in season.
 (They're) fresh. [They're] fresh. Of course [they're] fresh. [They're] in season now.²⁸

Thus, in example (1) the tone-units *e²-dan⁴* and *zhu³-tang¹* constitute a tone-group, functioning as the subject of the utterance, while the tone-units of *bu²* and *tai⁴-hao³*, also forming a tone-group, constitute its predicate. In (2) the tone-unit *ran²-hou⁴* itself constitutes a tone-group, functioning as an adverbial phrase, the tone-units *duo¹ shao^{3/2}* and *qian²* form a tone-group which serves as the predicate, and the subject of this utterance is formed by the tone-units *zhe⁴-ge⁰* and *lu³-rou⁴*, which also form one tone-group²⁹. In sentence (3) the tone-group formed by the tone-units *ni³* and *shuo¹ de⁰* functions as the subject, the tone-group formed by the tone-units *zuo⁴ hao³ de⁰* and *man²-tou²* serve as the subject complement, while *shi⁴*, the predicate of the utterance, is both a tone-unit and a tone-group. Finally, utterance (4) contains 4 major intonation-groups, the first two consisting of only a predicate and its implicit subject, which is known to the conversationalists from the preceding discourse context. In both cases this predicate is formed by one tone-group, consisting of the single tone-unit *xin¹-xian¹*. Much the same applies in the third sentence, which however, also contains the tone-group, and tone-unit, *dang¹-ran²*, serving as a disjunct. In the fourth sentence, the tone-unit *xian⁴-zai⁴* also constitutes a tone-group, serving as an adverbial of time, while the tone-unit, and tone-group, *dang¹-ji⁴* functions as the predicate, whose subject is again implicit. As the next point, we now come to the role of length, loudness, pitch and pause in the signaling of the boundaries of these units of *tonality*.

3.2.2 The signaling of intonation-group, tone-group and tone-unit boundaries

The signaling of the (major and minor) intonation-group boundaries and the boundaries delimiting tone-groups and tone-units touches upon the phenomenon of *declination*, or *down-trend*, to use the generic term. Therefore, we will first take a brief look at how this phenomenon is manifested in Mandarin Chinese and how it may be of relevance to *tonality*.

Declination in Chinese

It has long been known that Chinese utterances are affected by pitch-declination, described as the lowering of the (initial) pitch of consecutive tones within a stretch of speech (Chao 1968, Norman 1988, Tao 1996). Chao, for example, notes that ‘in [Chinese utterances] there is slight tendency for the pitch to trail off to a lower key towards the end’ (1968:40). Following Norman, declination in Chinese is defined here as the phenomenon that the pitch of a given tone towards the end of an utterance is generally lower than the pitch of the same tone spoken at the beginning (1988:149). Rather than lowering the pitch of single syllables, however, as is the case in non-tone languages, declination in Chinese affects the entire functional domain of the tones, the *tonal band*, which is moved downwards as a whole, together with its five levels³⁰. This is shown schematically in Fig. 17 below, in which a speaker message containing

²⁸ In sentence (2) the tone-units *zhe⁴-ge⁰* and *lu³-rou⁴* may also be seen as constituting a tone-group of their own, in which case the first serves as a premodifier of the second.

²⁹ Note that the tone-unit *zuo⁴ hao³ de⁰* and *man²-tou²* can also be analysed differently, in which case both constitute tone-groups of their own, the first serving as the premodifier of the second.

³⁰ This definition clearly implies that calculations of Chinese declination must be based on syllables with the same phonological status, i.e. bearing the same lexical tone, or at least occupying the same position within the tonal band. As a second prerequisite, given the above

three intonation-groups and representing the speaker message ‘Wo³ de^o mao⁴-zi^o! Wo³ de^o mao⁴-zi^o! Hai² mei² zhao³-dao⁴ ah^o!’³¹ is plotted on a slowly descending *tonal-band*, the lexical tones represented as strokes - for stressed, non-reduced tones - and *dots* - for neutral/neutralized tones.³² As Fig.17 shows, declination in Mandarin Chinese, gradually shifts the entire tonal-band downwards, affecting all the lexical tones contained within the stretch of speech.

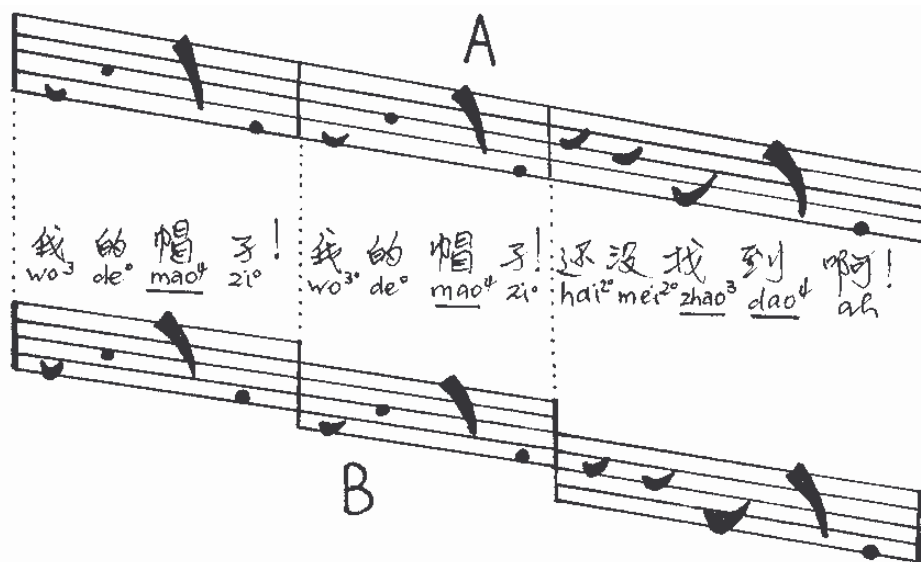


Fig. 17 Declination in Chinese (1)

Though the existence of declination in Mandarin Chinese speech has been confirmed by recent studies in instrumental phonetics, no consensus has yet been reached as to whether it is to be regarded as a phonological process or merely a phonetic phenomenon. While some researchers, such as Gårding (1987) and Shen (1990), have suggested that degree of pitch-declination in Chinese is linked to the syntactic or illocutionary force of the corresponding utterance, others, such as Tsao (1967) and Li (2002), see it as a purely phonetic process devoid of communicative function, and dependent, among other things, on the length of utterances and the amount and degree of stressing contained therein.

As a small contribution to the discussion around this phenomenon, an examination was conducted of the declination present in some of the Chinese speech samples recorded for this work. While the results of this examination confirmed that different degrees of declination are found in Chinese utterances³³, the obtained data also suggest that the degree of pitch-declination - or perhaps more objectively, of pitch-lowering - may vary across one and the same utterance. This is apparent in Fig. 18, showing the Chinese speech sample *yLong 2*³⁴: Thus, while *bottom-line* declination is gradual and steady during the first half of the utterance, descending from *nǐ*³ (228 Hz) to *mǎi*³ (220 Hz), a sudden, much stronger lowering of pitch occurs towards the end, between *mǎi*³ (220 Hz) and *lǒng*³ (156 Hz)³⁵. A similar strong shift is

discussed influence of prominence/stress on the pitch-height of Chinese tones, only the pitch/frequencies of syllables bearing the same degree of prominence/stress may be used for calculations of Chinese declination.

³¹ The English translation of this sentence, describable as an impatient exclamation, is ‘My hat! My hat! [I] haven’t found [it] yet!’

³² The difference between the two representations A and B will be discussed below.

³³ Declination was calculated for these utterances in the manner explained in Chapter 8. More will be said of the role of declination in Chinese in Chapter 9.

³⁴ The labeling system of the Chinese and German speech samples is explained in Chapters 6 and 7. A representation of these utterances in Chinese script and *pin-yin* is found in Section 2 of Appendix I, where translations are also given. The *pin-yin* transliteration system is explained in Section 1 of Appendix II.

³⁵ Down- and upshifts are represented with the symbols ▴ and ▾, respectively.

pitch is also found to Fig. 19, showing *yMan 2*: Here, the pitch/frequency difference separating *ni*³ (159Hz) and *mai*³ (114 Hz) is very large, suggesting a sudden increase in *bottom-line* declination, or alternatively, a sudden *downshift* in pitch.

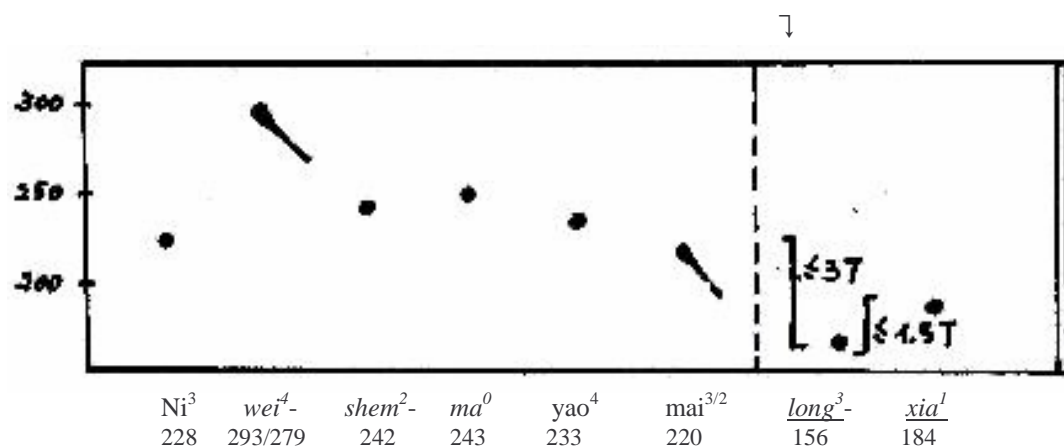


Fig. 18 Declination in Chinese (2) (*yLong 2*)³⁶

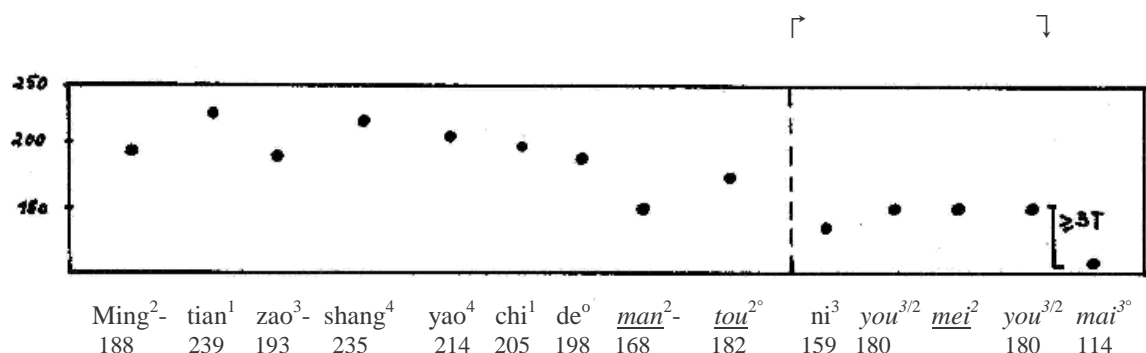


Fig. 19 Declination in Chinese (3) (*yMan 2*)

Finally, both *down-* and *upshifts* in pitch of the *bottom-line*, the latter a sudden *rise* in pitch/frequency, are found in *yMing 1* and *wMan 3*, shown in Figs. 20 and 21. Thus, in Fig. 20 there is an *upshift* on *ni*³ (222 Hz), which is 5 Hz higher than the preceding *jie*³ (217 Hz), and a *downshift* on *ming*² (185 Hz), 10 Hz lower than the preceding *dian*³ (196 Hz). In Fig. 21, there is clearly an *upshift* on *ye*³ (235 Hz) which is 22 Hz higher than the preceding *man*² (213 Hz), which is followed by a *downshift* on the second *man*² (182 Hz), 53 Hz lower than the preceding *ye*² (235 Hz). Though nothing definite can yet be said here about the precise locations or communicative functions of such *up-* and *downshifts*, Figs. 18 to 21 suggest that these strong and sudden movements in pitch/frequency are more frequent at intonation-group and tone-group boundaries than between the tone-units of one tone-group. Indeed, the absence of *downshifts* between and especially within tone-units would seem in the interest of communication as the tonal structures of these morphemic units would thus remain intact. On the other hand, it seems unlikely that the long stretches of speech which are often covered by sequences of tone-units should remain unaffected by declination, so that one would certainly expect the presence of gentle, gradual declination within tone-groups. In Fig. 17, the bottom half (B)

³⁶ These representations are based on the *f0*-contours of the Chinese utterances, obtained in the manner described in Chapter 8 (Section 8.2.1). The tones of syllables of long duration appear as 'tadpoles', while dots mark syllables which are short and whose tones are mostly reduced or neutralized. Note that the relative length of syllables as shown in *f0*-contours not only reflects stress/prominence, but is also dependent on speech tempo. Therefore, both 'tadpoles' and 'dots' may represent prominent syllables. There is no indication of relative intensity in *f0*-contours.

of the representation of declination in Chinese includes the pitch-shifts discussed here, in this case limited to the intonation-group boundaries³⁷. The last 2 examples to be discussed here are

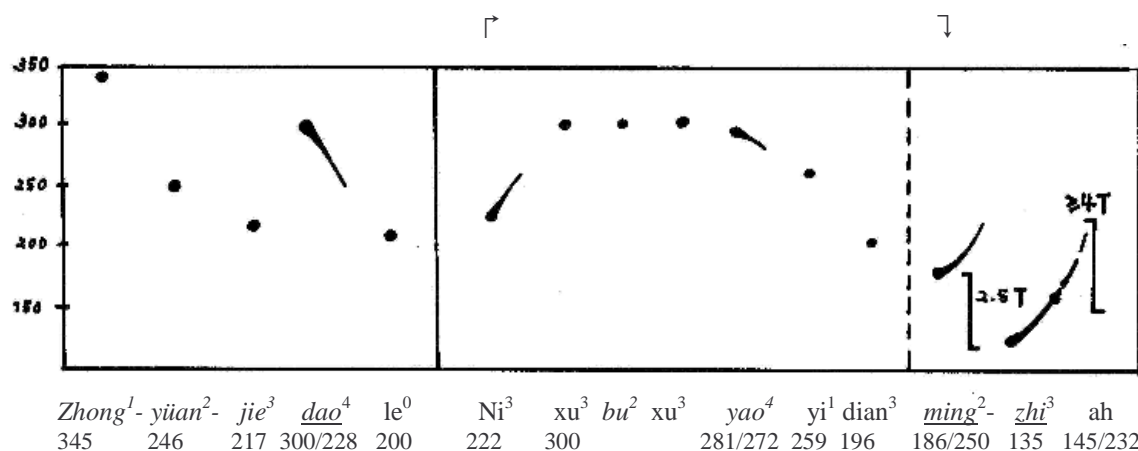


Fig. 20 Declination in Chinese (4) (yMing 1)

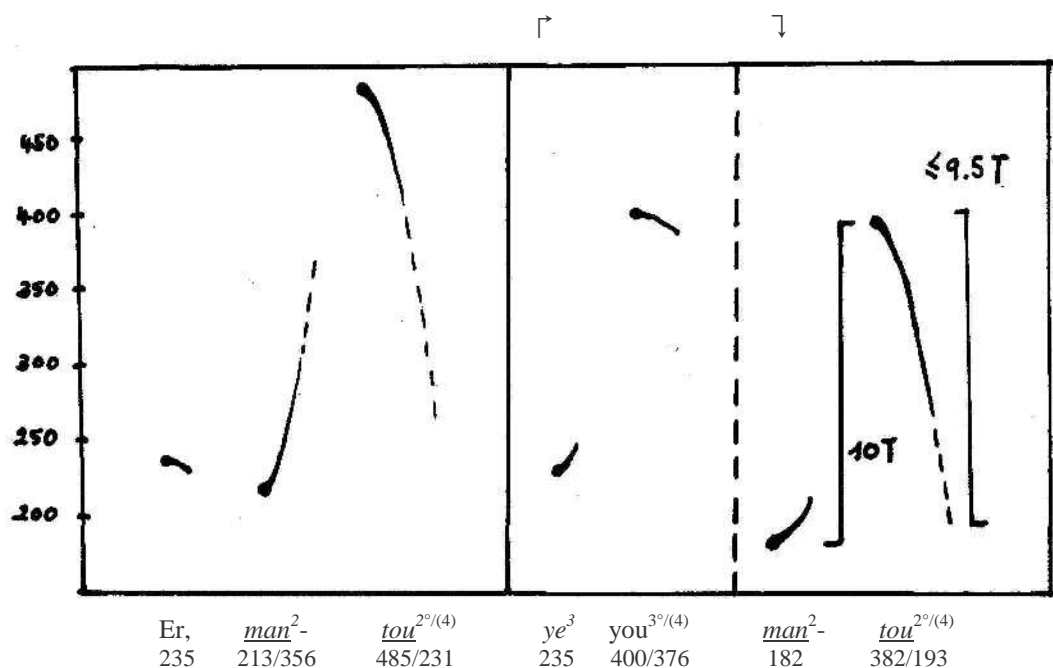


Fig. 21 Declination in Chinese (5) (wMan 3)

wNing 4 and *yLa 2*, shown in Figs. 22 and 23. They suggest that intonation-group boundaries in Chinese are not always marked by *up-* or *downshifts*: In *wNing 4* there is no evidence of a *downshift* at the minor intonation-group boundary³⁸ between *meng*² and *pao*⁴, and the same applies in *yLa 2* to the major boundaries between *ne*⁰ and *wo*³ and between *jiao*¹ and *ni*^{3/2}. Interestingly, the *downshifts* that do occur in Fig. 23 (on *mai*³) both occur after a pause, suggesting that at least in this utterance, we are dealing with a phonetic phenomenon³⁹.

³⁷ In the upper half (A) of Fig. 17 the same utterance is represented without these pitch-shifts.

³⁸ The stretch of speech *pao*⁴ *cha*² *he*¹ is seen here as representing an adverbial clause of intent.

³⁹ Much the same was observed on marked changes in pitch at intonation-group boundaries in English and German. See Section 2.1.1.3.

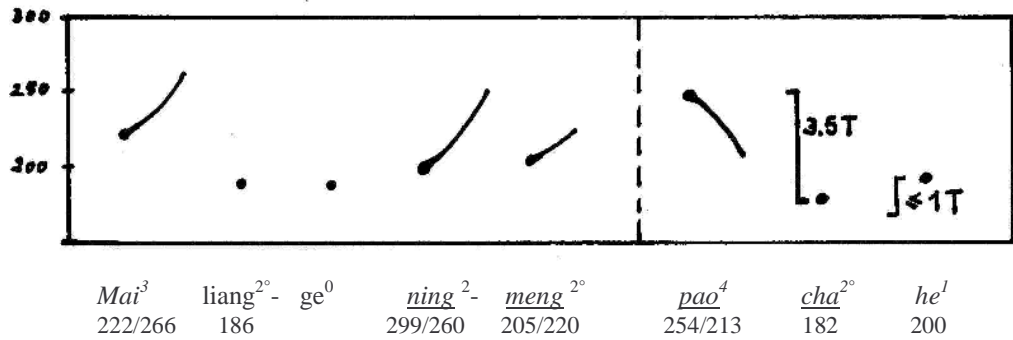


Fig 22 Declination in Chinese (6) (wNing 4)

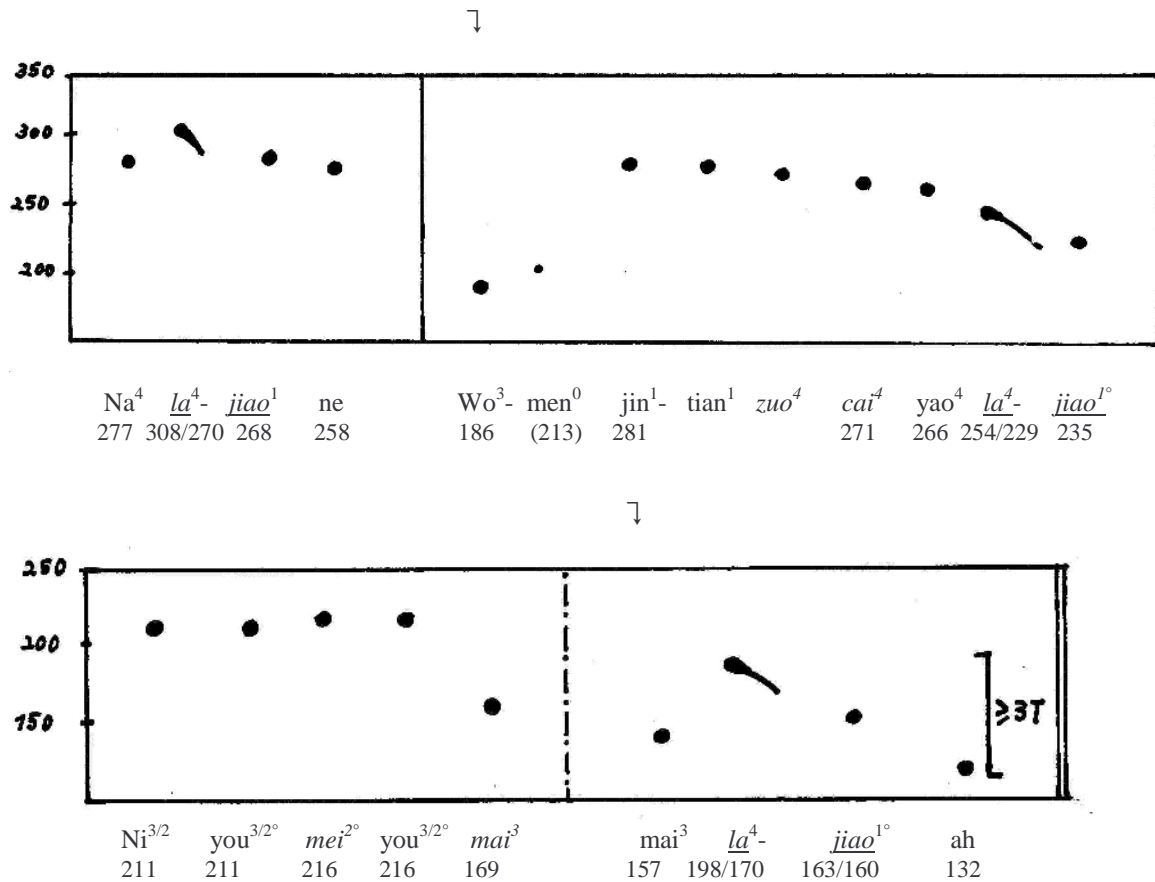


Fig. 23 Declination in Chinese (7) (yLa 2)

To conclude the present discussion, the speech data discussed here suggest that *down-* and *upshifts*, defined as sudden, relatively strong and thus readily perceivable down- and upward shifts in pitch/frequency, only have a limited communicative function with respect to *tonality*. While they often occur at intonation- and tone-group boundaries, the fact that they are not always present at such points seems to rule out the possibility of treating them as markers of these boundaries. On the other hand, the above discussed data indicate that *up-* and *downshifts* do not normally occur within or between tone-units, unless after hesitational pauses, and that only declination may be expected within tone-groups. We now return to the main topic of this subsection and summarize the points which may be concluded from the above discussion with respect to the signaling of intonation-group, tone-group and tone-unit boundaries:

(1) *The signaling of (major and minor) intonation-group boundaries*

Like their English/German correlates, major intonation-group boundaries in Chinese are generally marked by pauses of varying length preceded by pre-boundary slow-down (Chao 1968, Tao 1996, Yang & Wang 2002), while minor boundaries are preceded by slow-down only (Yang & Wang 2002, Tao 1996). The results of the above discussion suggest that *up-* and *downshifts* of the *bottom-line* are sometimes, but not always, found at these boundaries.

(2) *The signaling of tone-group boundaries*

Whether tone-group boundaries are accompanied by preboundary slow-down or pauses cannot be said with certainty at this point. However, given the fact that these units represent clause constituents, such as subjects, topics, objects and predicates, the presence of phonological - i.e. non-hesitational - pauses is unlikely here⁴⁰. The Chinese speech data discussed above suggest that *up-* and *downshifts* may, but again need not necessarily, accompany tone-group boundaries.

(3) *The signaling of tone-unit boundaries*

As was observed above, it would seem in the interest of communication if tone-units were affected only by declination, which the human perceptual apparatus compensates for. For the same reason, the absence of pauses and preboundary slow-down at tone-unit boundaries seems more likely than their presence⁴¹.

3.3 *Tonicity in Mandarin Chinese*

Chinese tonicity, involving the signaling and marking of the *focus*⁴², a process also referred to as *focusing* (Kratochvil 1998), is an area of strong controversy, part of which seems to be due to the fact that there is no agreement on how to locate focused syllables, and how to avoid confusion with phonologically prominent syllables of morphemic compounds⁴³. According to an early definition by Chao (1968), the general rule is that the *focus* falls on the last stressed syllable of an utterance (1968: 42). However, as long as the precise manner in which the Chinese *focus* is signaled by length, loudness, pitch and pause, remains unclear, the validity of this proposition cannot be examined⁴⁴. While it is generally agreed that the primary marker of *focusing* is an enlargement of the tonal band-width, i.e. a widening of the pitch-range covered by the affected syllable, the precise degree(s) of enlargement involved have yet to be established. It is also still largely unclear how this enlargement of tonal band-width compares to the widening of the tonal band-width encountered in phonologically prominent syllables. As a third point, there is also disagreement with respect to the role of duration and intensity: While Chao (1968), Norman (1988) and Kratochvil (1998) see increases in duration and intensity as markers of the *focus*, Shen (1990) excludes intensity⁴⁵. Finally, it has not been definitely clarified if the increases in tonal band-width encountered in focusing also incur changes in the initial pitch of tones, as observed in prominent syllables by Kratochvil (1998: 422)⁴⁶.

⁴⁰ Tao 1996 arrives at much the same conclusion (1996: 35ff.). According to Cruttenden, non-hesitational, i.e. phonological, pauses are not found between these syntactic units in English (1986: 37).

⁴¹ Interestingly, the evidence discussed in Yang & Wang 2002 does suggest the presence of pauses as well as preboundary slow-down at these points. This could be linked to the fact that these researchers examined read speech. This assumption is supported by the fact that Klatt (1975), who reports similar findings for English, is also relating to read, not spontaneous, speech.

⁴² The broad definition of the *focus* as falling on the most important item of information within one intonation-group is adopted from the discussion of English/German intonation in Section 2.2.1.1. The term *nucleus* is avoided here, since at this stage it cannot be ascertained whether the *nucleus* of western intonation and the *focus* of Chinese intonation are indeed formally and functionally equal.

⁴³ Naturally, as long as the identification of the Chinese *focus* is a problem, an examination of the rules of Chinese *tonicity*, relating to the position of broad and narrow *focus* in Chinese utterances is not possible.

⁴⁴ As in English and German, it is to be expected that the position of the Chinese *focus* is also dependent on syntactic and semantic factors of the kind discussed above in Section 2.2.1.1.

⁴⁵ See discussions in Chao 1968: 35, Norman 1988: 148, Kratochvil 1998: 426f. and Shen 1990: 59 ff.

⁴⁶ As prominence and focusing both involve enlargements of tonal band-width, it is to be expected that the same changes to initial pitch of

As a contribution to the discussion on Chinese *tonicity*, a first analysis of the utterances in the Chinese speech corpus used in this work yielded two findings, shown in Figs. 24 and 25. As a first point, the fact that the Chinese *focus* is often not limited to one syllable but spreads over a number of consecutive morphemes (Kratochvil 1998), is clearly to be seen in both examples. Thus, in Fig. 24 (representing *yLong 1*), the *focus* is slowly built up on *cai*² *shi*²- and culminates on *-er*⁴ and in Fig. 25 (*wMan 2*) the *focus* begins with *pa*⁴ and ends with *jing*⁴. As a second point, a close examination of the pitch/frequency-range of these and a number of other Chinese utterances in the corpus showed that, with respect to pitch/frequency-range, it is difficult to differentiate between the *focus* of Chinese utterances and what might be referred to as the ‘stressed’ syllables of tone-groups. While in *wMan 2*, *man*² and *-tou*^{2°}, the stressed syllables of the tone-group from *ni*³ to *tou*^{2°}, are slightly narrower in pitch/frequency-range than *jing*⁴, the syllable which according to Chao’s rule represents the *focus*, on the other hand, in *yLong 1* the pitch/frequency-ranges of *long*¹-*xia*¹, the stressed syllables of the tone-group between *long*³ to *jin*¹ - are virtually identical to those of the focused syllables, *shi*²-*er*⁴.

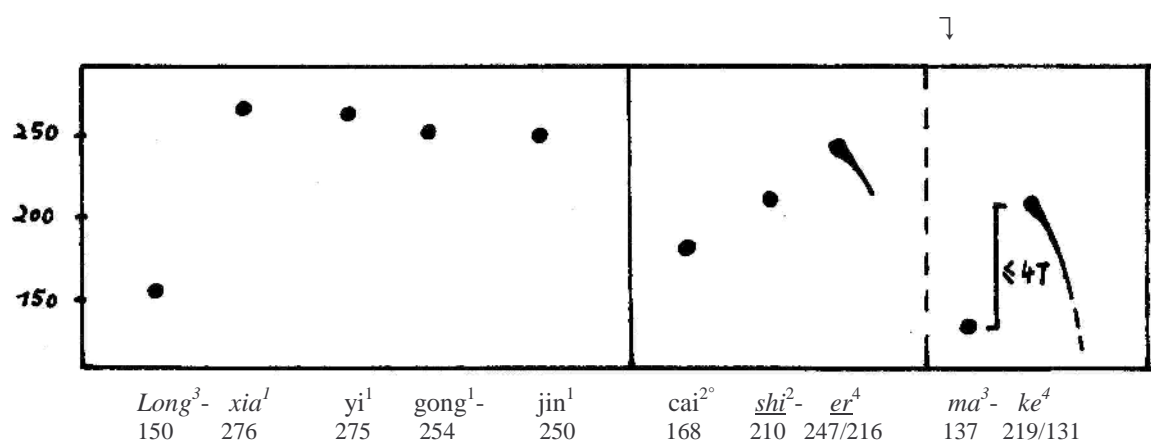


Fig 24 Mandarin Chinese tonicity (1) (*yLong 1*)

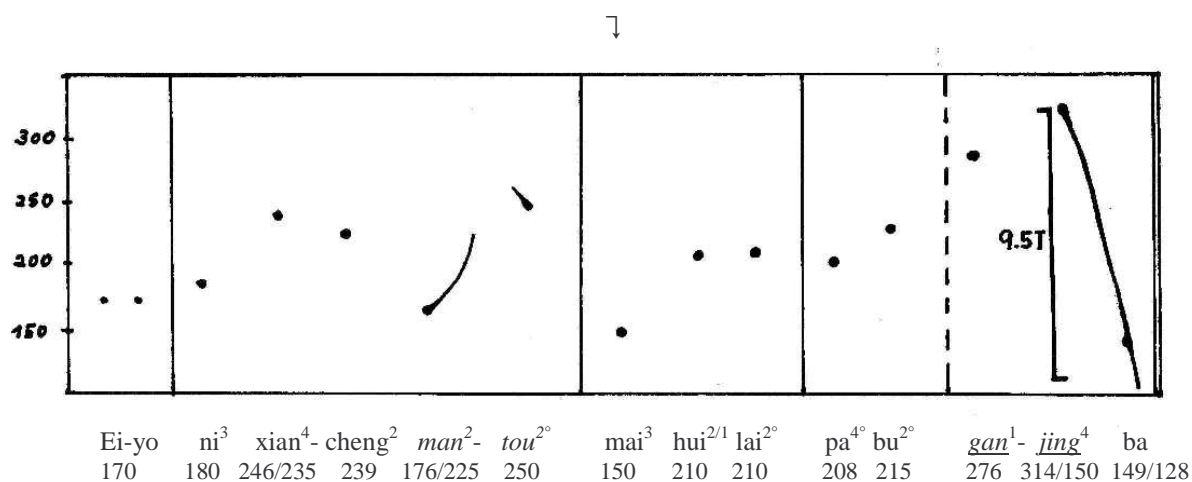


Fig. 25 Mandarin Chinese tonicity (2) (*wMan 2*)

tones occur in focusing as were described for prominent syllables above, where the initial pitches of 1st, 2nd and 4th tones are raised, those of 3rd tones lowered. Note that in order to differentiate between local rises in pitch and pitch-range, the latter involving a widening of tonal band-width, examinations of focusing should be based on more than one lexical tone. These should include one tone whose pitch-pattern is associable with the *top line* - ideally either a 1st or a 4th tone - and a second tone which may be associated with the *bottom line*, ideally a 3rd tone.

3.4 Intonational pitch-patterns in Mandarin Chinese

The possible existence of intonational pitch-patterns in Mandarin Chinese, i.e. pitch-movements functioning independently of the lexical tones and signaling communicative meaning of the kind discussed in the introduction to Part Two, is a highly controversial issue. Judging from the discussions of this topic in the literature, the general view seems to be that the phonological function of pitch in Chinese, in the form of the lexical tones, rules out all possibility of a non-lexical, intonational, use of pitch in this language. Despite early evidence to the contrary presented by Chao (1968), which is now finding confirmation in the results of recent studies in Chinese experimental phonetics, many scholars even today maintain that Chinese intonation cannot possibly involve more than variations in overall pitch, affecting whole clauses or sentences at a time, since local pitch-variations would distort the shape of the lexical tones (Shen 1990). In particular, the following three ‘basic principles’ are widely seen as unassailable: (1) In the interest of communication, the tonal structure of all morphemic units (*tone-units*) must remain intact, (2) Intonational pitch-patterns, in so far as they exist at all, may not occur independently of segments, but must always be bound to segmental form, and (3) Such intonational pitch-patterns may only occur at the very end of utterances. In the following we first discuss the conventional approach to Chinese intonation. Then we will turn to the evidence for intonational pitch-patterns in Chinese.

3.4.1 Utterance-final intonational pitch-patterns on sentence-final particles

Presumably due to the ‘basic principles’ 2 and 3 given above, most traditional intonation research on the Chinese language has focused on utterance-final sentence particles⁴⁷. Kratochvil (1998), who prefers to call them ‘segmental carriers of intonation’, describes these as follows:

‘Intonation carriers, mostly but not necessarily monosyllabic items constituting a small closed set in [Mandarin Chinese], correspond to interjections and what are vaguely referred to as ‘sentence particles’ in Chinese linguistic terminologies.[...] Their most typical position is at the end of a sentence. Since they carry no tone, they are a subtype of atonic [i.e. neutral] syllables, and their prosodic properties are directly determined by intonation’ (1998: 426).

In his discussion of the communicative functions of such intonation carriers, Kratochvil contrasts the two sentences ‘*Ni zemma bu fu-shu ah?*’ and ‘*Ni zemma bu fu-shu ne?*’⁴⁸, the first an ‘irritated exclamation’, the second an ‘impatient question’. However, although he is nominally discussing Chinese intonation phenomena, there is no further mention of the term ‘intonation’. Instead, the discussion focuses on the choice of the particles, whose contribution Kratochvil obviously sees as primary, all the more as the two particles he is discussing here ‘carry the neutral tone’⁴⁹. In much the same manner, the vast majority of discussions of the intonation associated with sentence particles found in the literature focus exclusively on the ‘syntactic’ or ‘illocutionary’ functions of the particles themselves and disregard the contribution of the pitch-patterns they carry, or initiate.

While the meaning of a Chinese utterance certainly depends to some degree on the choice of the final particle, an exclusive orientation towards their segmental form is too narrow, since it ignores the fact that the pitch-patterns carried by the particles also contribute to their communicative function. This rather limited traditional view of Chinese intonation is presumably linked to the general view, also encountered in Kratochvil’s discussion above, that sentence-final particles possess one single, unchanging tone, usually the *neutral tone*, whose communicative role is negligible. A point to which no attention has been paid so far is the fact that the

⁴⁷ The particles most often discussed in the literature are *ma*, *ne* and *ba*, signaling different kinds of interrogative meaning, and (y)ah and oh, used for a larger range of meanings (Chao 1968: 795). In the second group, Chao also includes *le*, *de* and *me*, the latter of which he sees as a low allophone of *ma*. Not counted as particles are interjections, such as *ei*, *eijo* and *mmh*, encountered at the beginning or within utterances (1968: 815 ff.). Discussions of Mandarin Chinese particles, in particular *ma* and *ba*, are found in Han 1991, 1995. See also Chan <http://deall.ohio-state.edu/chan.a/mTobL.htm>

⁴⁸ The first sentence is translated by Kratochvil as ‘How come you don’t admit defeat?’, the second as ‘Why don’t you admit defeat?’

⁴⁹ More will be said of this position below.

tones, or better the *pitch-patterns*, carried by the majority of Mandarin Chinese particles are in fact not fixed and unchanging, but vary in a subtle manner from one utterance to the next, signaling meanings which are often infinitively more subtle than could be described in terms of syntactic or illocutionary function⁵⁰. This applies especially to the particles *ba*, (y)*ah* and *oh* which, strictly speaking, have no fixed tone. Variations in pitch-patterns are, however, also found with *le* and *de*, carrying the neutral tone, and with the interrogative particles *ma*, *ne*, whose pitch-patterns are drawn out and relatively high in pitch, resembling 1st tones. The variability of the pitch-patterns carried, or better *initiated*, by final particles is illustrated in the two utterances shown in Figs. 26 and 27, uttered by the same native Chinese speaker. In each utterance, the sentence particle (*ba*) is used with a different pitch-pattern - a steep and deep *fall* in *wMan 2* and a gradual and gentle slope-like fall in *wLu 4*. The fact that these two utterances, differing only in the pitch-patterns carried by their common final particle, were judged by native Chinese speakers to signal different speaker attitudes, broadly describable as 'strong' and 'insistent' in the first case, 'gentle' and 'deferent' in the second, indicates that these two pitch-patterns may indeed be regarded as communicating different meanings.

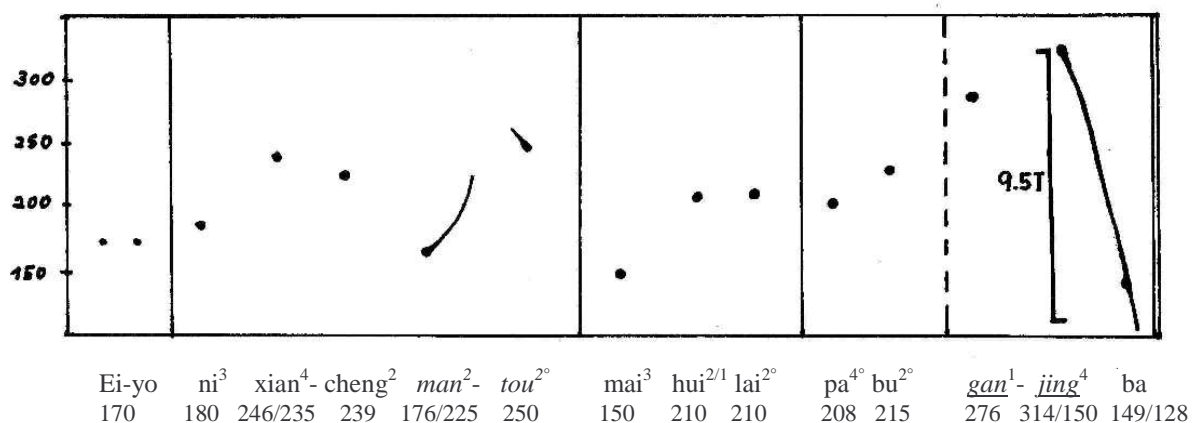


Fig. 26 The communicative functions of Chinese particles (1) (*wMan 2*)

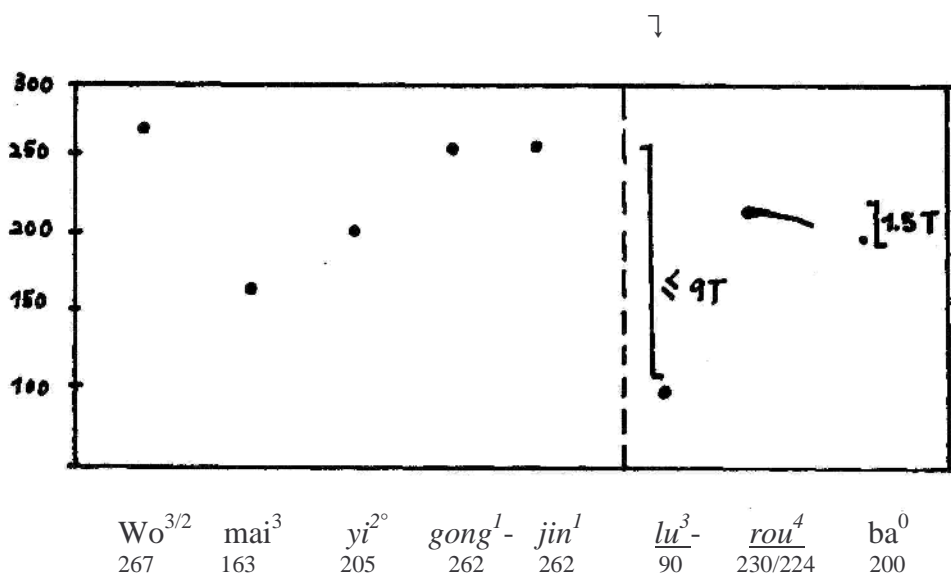


Fig. 27 The communicative functions of Mandarin particles (2) (*wLu 4*)

⁵⁰ The meanings signaled by such variations are, however, generally within the semantic scope of the sentence particles themselves.

The two following examples, shown in Figs. 28 and 29, illustrate the fact the the particle *ah* may also carry different pitch-patterns.

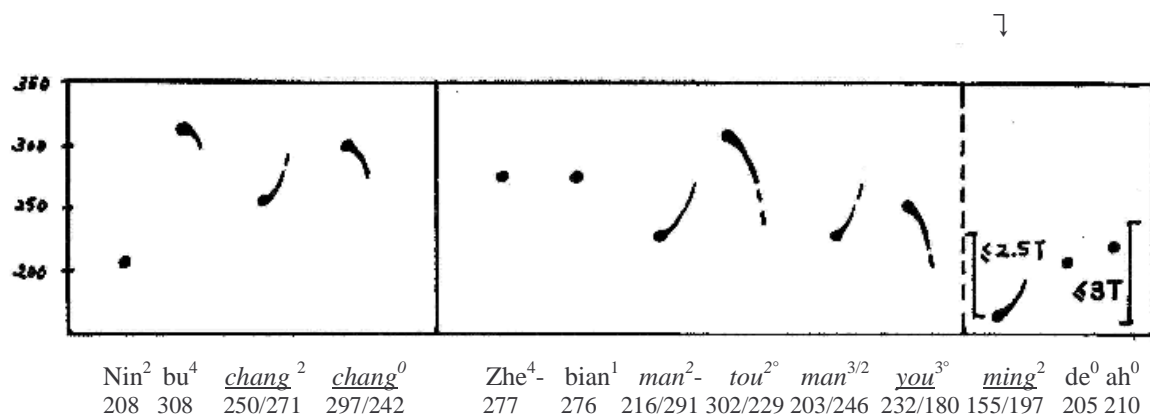


Fig. 28 The communicative functions of Mandarin particles (3) (*yMan 1*)

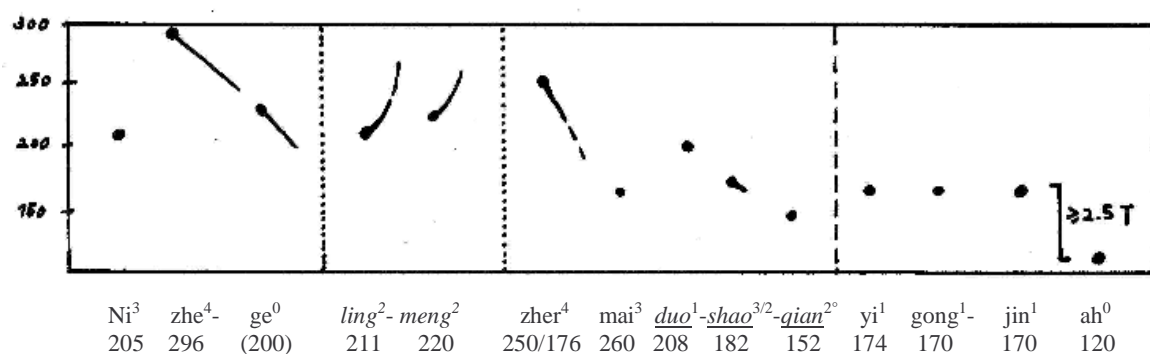


Fig. 29 The communicative functions of Mandarin Chinese particles (4) (*yLing 2*)

The two utterances shown in Figs. 28 and 29 were judged by native Chinese speakers to signal a strong and rather direct attitude on the part of the speaker. In addition to this, the second utterance, in which *ah* carries a low-rising pitch-pattern, was seen to communicate a challenging, possibly menacing, attitude. These examples suggest that far from merely accompanying their segmental form, the pitch-patterns of Mandarin Chinese particles function as carriers of intonational meaning.

3.4.2 Utterance-final intonational pitch-patterns without sentence particles

Chao (1968) was the first linguist to suggest that intonational pitch-patterns may also be realized in Mandarin Chinese *without* final particles. He describes what he refers to as falling and rising ‘intonational endings’ as follows:

These two intonational endings [...] are of a very special morphophonemic nature. [What is special] about each of these two particles is that it does not have a segmental phoneme of its own but resides parasitically on the last morpheme by prolonging it for the length of a neutral-tone syllable on which to put a rising or falling ending. (1968: 812)

Among the examples Chao gives in support of this theory are jin⁴ lun¹-dun^{1↑}? (‘near London?’), in which the 1st tone of dun¹ is ‘extended’ so to say, by an additional *rise*, and you³ jiu^{3↓} (‘There is wine’), where a falling pitch-pattern is ‘added onto’ the 3rd tone of jiu³ (1968: 812). No final particles, are added in either utterance to function as carriers of these pitch-patterns. Among the meanings suggested by Chao for these two intonation patterns are:

‘incredulity’, ‘impatience’ and ‘insistence’ for the *rising pitch-pattern*, and ‘liveliness’, ‘heaviness’, ‘remonstrance’ and ‘pretended (in the sense of exaggerated) emotion’ for the *fall*⁵¹. Chao’s theory of falling/rising non-segmental intonation pitch-patterns in Chinese sparked a fierce debate and was rejected by most researchers in the field on the grounds that such additional pitch-patterns would distort the tones of the affected morphemes beyond recognition⁵². However, the findings of recent instrumental speech analyses indicate that these phenomena do in fact exist⁵³. One very conspicuous intonational pitch-pattern of precisely this kind is also found in one of the utterances in the Chinese speech corpus recorded for this work and shown in Fig. 30 (*wNing* 2). Here, the final syllable of *ning*²-*meng*² carries an additional *falling* pitch-pattern, which is added onto that of its 2nd tone. This utterance was judged by native speakers in terms of a rather exaggerated emotionality on the part of the speaker, this being one of the meanings suggested for this pitch-pattern by Chao⁵⁴.

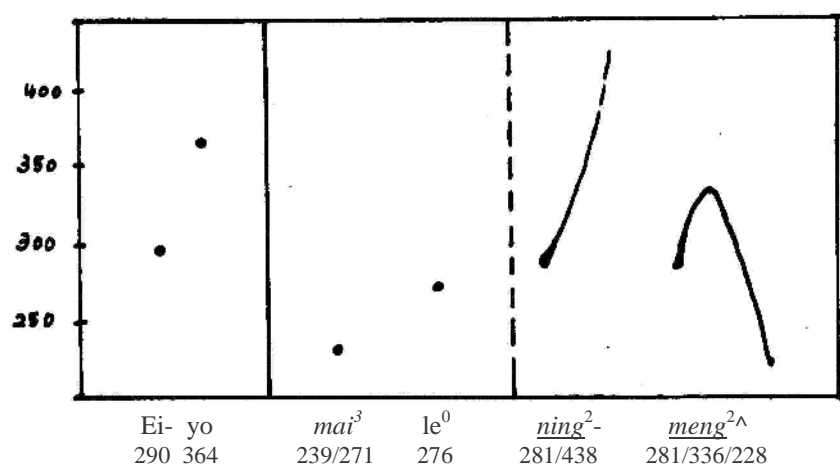


Fig. 30 Intonation in Mandarin Chinese without segmental particles (*wNing* 2)

3.4.3 Intonational pitch-patterns across Mandarin Chinese utterances

The two kinds of intonational pitch-patterns in Mandarin Chinese discussed up to this point were limited to single syllables found at the very end of utterances. A third kind is suggested here based on the writer’s observation that certain (short) sequences of utterance-final syllables can also carry or, more correctly *trace*, intonational pitch-patterns. However, as the addition of *falls* or *raises* onto the tones of non-final syllables would certainly interfere with communication by distorting the tonal structures in within the effected morphemic units, it is suggested that only the initial pitch-height of these syllables, and thus, in effect, their *pitch-range* (tonal band-width), is affected. Thus, the proposal here is that the (initial) pitch-heights of such sequences of utterance-final syllables create what could be referred to as intonational pitch-patterns, or contours, signaling intonational meaning of the kind discussed for English and German intonational pitch-patterns in Chapter 2. The creation, or tracing, of such utterance-final intonational pitch-patterns will be referred to here as the ‘final tonal behavior’ of utterances. Though nothing precise may yet be said here with regards to the location of these pitch-patterns, Figs. 31 - 36 may be regarded as evidence for their existence and also suggest that they are linked to the last one or two tone-units and/or the last stressed syllable of an utterance, most often its *focus*⁵⁵.

⁵¹ Chao 1968: 813f. It will be noted that these meanings are highly reminiscent of the attitudinal functions assigned to the nuclear pitch-patterns of English/German intonation (See Chapter 2).

⁵² See the detailed discussion in Shen 1990, according to whom Chao’s theory was supported by Rumjancev (1972), but opposed, among others, by Chang (1958) and Ho (1977).

⁵³ A number of examples of falling pitch-patterns clearly of the kind which Chao envisaged are also found in Yang & Campbell 2001.

⁵⁴ More will be said of this pitch-pattern in Chapter 9.

⁵⁵ The communicative, and particular attitudinal, functions of all three types of intonational pitch-patterns will be examined in Chapter 9.

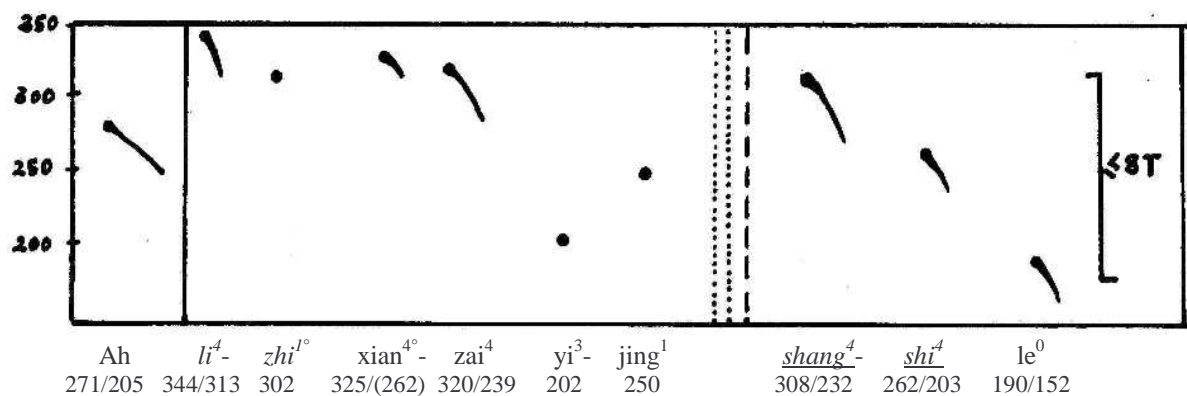


Fig.31 Intonational pitch-patterns across Chinese utterances (1) (yLizhi 4)

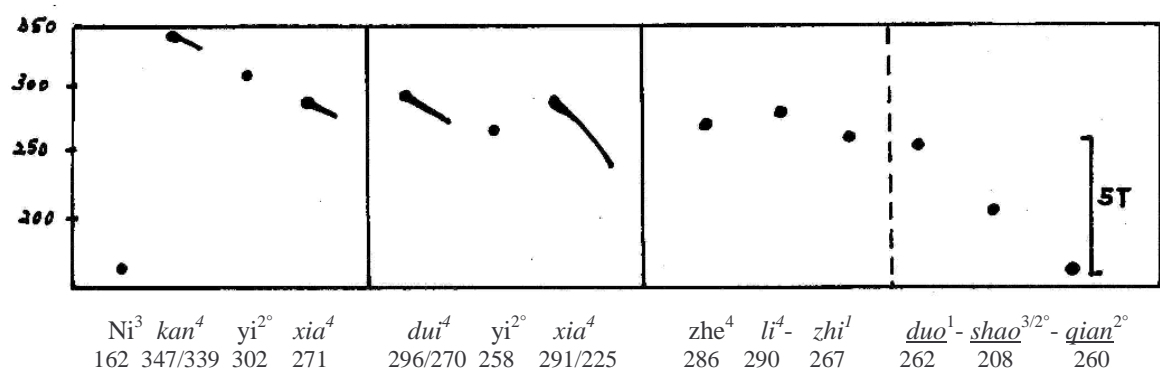


Fig. 32 Intonational pitch-patterns across Chinese utterances (2) (yLizhi 2)

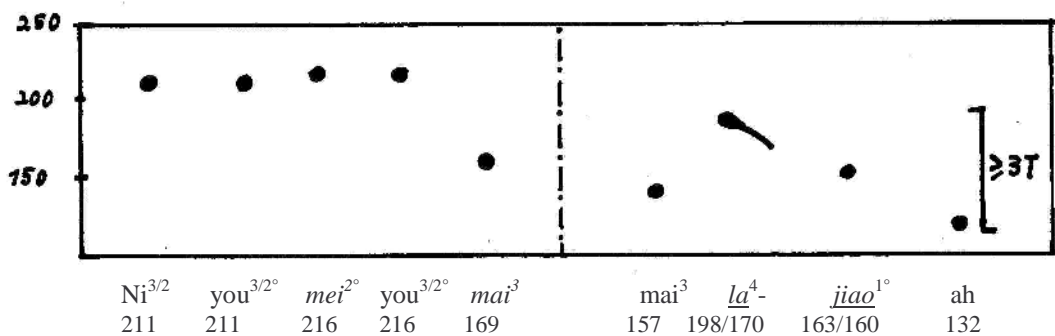
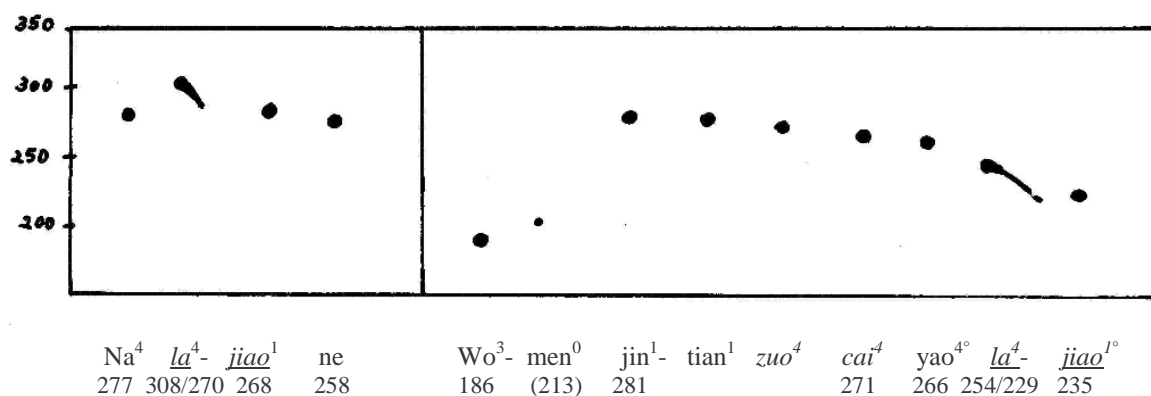


Fig. 33 Intonational pitch-patterns across Chinese utterances (3) (yLa 2)

In Figs. 31, 32 and 33, the (initial) pitch of the tones in the last tone-unit all trace relatively steeply falling patterns, reminiscent of the *wide falls* of English and German intonation. The pitch-patterns found in Figs. 34 and 35, on the other hand, represent abrupt, wide drops in pitch, followed by what could be described as the Mandarin Chinese equivalents of the *flat tails* of English intonation: a sequence of utterance-final tones, or syllables, all carrying approximately the same pitch/frequency. As a final point, as may be seen in the utterance pre-

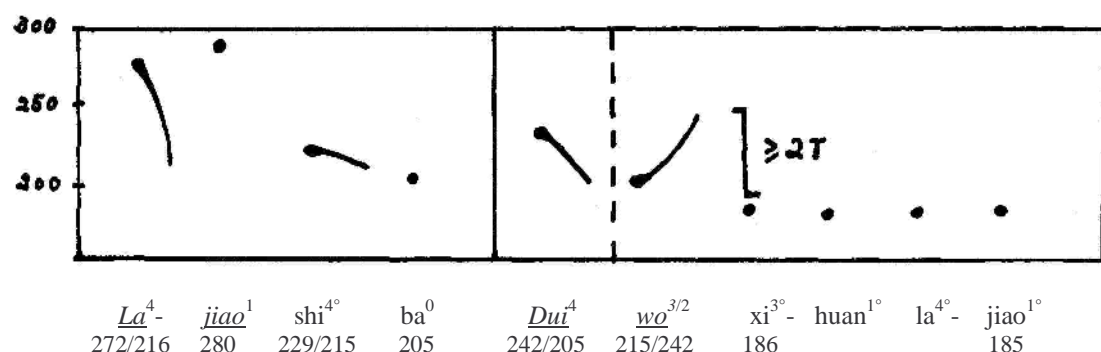


Fig. 34 Intonational pitch-patterns across Chinese utterances(5) (wLa 4)

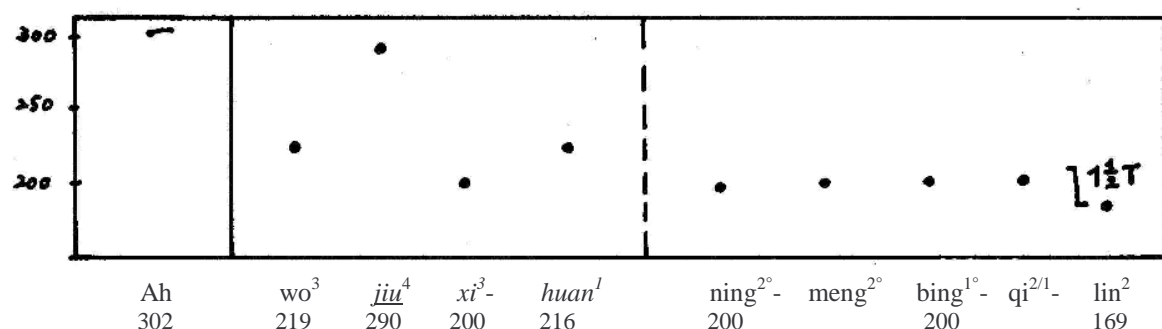


Fig. 35 Intonational pitch-patterns across Chinese utterances (6) (wBing 3)

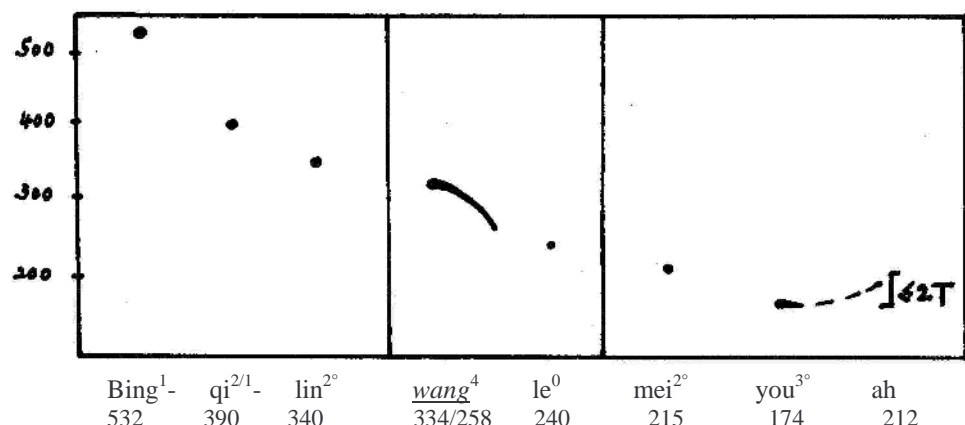


Fig. 36 Intonational pitch-patterns across Chinese utterances (4) (wBing 2)

sented in Fig. 36, some utterances contain two instances of 'final tonal behavior'. Thus, in Fig. 36 the pitch-pattern traced by the syllables *wang*⁴ and *le*⁰, corresponding to an abrupt and wide fall, is followed in the last tone-unit by what could be described as the Mandarin

Chinese equivalent of the *low rise* of English/German intonation. As the first instance of this tonal behavior (on *wang*⁴ and *le*⁰) is associated with the *focus* of the utterance, *wang*⁴, this first type of tonal behavior will be referred to here as *focal tonal behavior*. The second instance of tonal behavior, falling on the utterance-final syllable, will be referred to as *caudal tonal behavior*. While nothing precise may yet be said on their functions, in analogy to what is known for western intonation, it is suggested here that *focal tonal behavior* generally plays a more important, *primary*, role in the signaling of meaning, than *caudal tonal behavior*. This issue will be examined in Part Three⁵⁶.

3.5 Summary: The communicative forms of Mandarin Chinese intonation

To summarize the above discussion, based on the evidence in the literature and a first analysis of the Chinese speech samples recorded for this work, it may be concluded that Mandarin Chinese differentiates between the same three communicative forms as are found English and German, intonation: (1) *Tonality*, the division of connected speech into units of tonality; (2) *Tonicity*, the position of the *focus* within the corresponding utterance; (3) *Tone*, intonational pitch-patterns signaling communicative meaning of the kind defined above in the introduction to Part Two. On the other hand, the evidence discussed above suggests that the signaling, or creation, of these *communicative forms* is a little different here than in English and German: Though the fundamental role of length, loudness, pitch and pause in the signaling of tonality and tonicity, is the same, it seems that the role played by loudness in the signaling of the Chinese *focus* is less important than its role in the marking of the *nucleus* in English and German. On the other hand, *tonal band-width*, corresponding to the pitch/ frequency-range covered by the lexical tones over a given stretch of speech, is deeply involved in *focusing*. Coming now to the intonational pitch-patterns of Mandarin Chinese, referred to here as the (final) ‘tonal behavior’ of utterances, pitch-range/tonal band-width also appears to play a role in their creation. Evidence was found for three different kinds of intonational pitch-pattern⁵⁷ in Mandarin Chinese: (1) pitch-patterns carried by utterance-final sentence particles, (2) ‘non-segmental’ patterns ‘added onto’ the tones of the very last syllables of utterances, and (3) pitch-patterns which are traced by the initial pitch, or pitch-range, of certain sequences of utterance-final tones. While the precise location of the third type of final tonal behavior cannot be ascertained yet, the evidence discussed here suggests that it begins on the last strongly stressed syllable of an utterance, in general its *focus*, and extends up to the end of the utterance. The tonal behavior found on the *focused* syllable is referred to here as *focal tonal behavior*, while that often found at the end of the utterance is referred to as *caudal tonal behavior*. Finally, with respect to their communicative functions, it was noted that the - basically attitudinal - meanings discussed by Chao (1968) for his ‘added-on’ pitch-patterns, corresponding to type (2) above, are very similar to those associated with the nuclear pitch-patterns of western intonation. It is possible that similar meanings will be found to apply to the pitch-patterns of sentence final particles and final tonal behavior, a proposition to be examined in Part Three, Chapter 9⁵⁸.

⁵⁶ See in particular Chapters 9 and 10.

⁵⁷ Note that, given the fact that these are strictly speaking variations in pitch-range, it would be more accurate to speak of pitch-range patterns.

⁵⁸ Though this cannot be ascertained here yet, it is possible that what is referred to here as caudal tonal behavior, found at the end of focal tonal behavior (type 3 above) is formally identical to the pitch-patterns carried by sentence-final particles (type 1) and the pitch-patterns ‘added-on’ to non-particle final tones of utterances (type 2). This will also be examined in the experimental part (See Chapter 9).

4. The communicative functions of overall pitch, loudness, tempo and fluency

As was observed above in the introduction to Part Two, *overall pitch, loudness, tempo and fluency* constitute a second set of communicative forms based on pitch, loudness, tempo and pause. In contrast to the local variations in these features which give rise to the *forms of intonation*, described in Chapters 2 and 3, the *overall features* of pitch, loudness, tempo and fluency are based on longer-term variations in pitch, loudness and tempo and the presence of pause phenomena which extend over one or more utterances, or intonation-groups. Overall pitch, loudness, tempo and fluency is sometimes also described for longer stretches of speech, comprising one or more *paratones*¹. Like the forms of intonation, the overall forms can serve to signal a number of different communicative functions, such as the linguistic, paralinguistic and extralinguistic functions to be discussed below. Though research in this field has been mainly limited to the English language, these communicative processes have been found to originate in universal physiological processes, so that the communicative functions discussed below for English also apply to a certain degree to German and Chinese².

Before beginning with the discussion of their communicative functions, the overall features are presented and defined briefly as follows.

4.1 Definitions - overall pitch, loudness, tempo and fluency

4.1.1 Overall pitch³

The overall, or mean, pitch of an utterance, also referred to as its *register*, is defined as the mean pitch-height of its syllables, excluding the brief, local variations in pitch which contribute to the signaling of the forms of intonation. The overall pitch of an utterance is calculated by determining the frequency of each consecutive syllable and determining their mean value, indicated in Hertz (Hz.).

4.1.2 Overall loudness

The overall, or mean, loudness of an utterance is defined as the mean loudness of all the syllables contained therein, excluding the short-term variations in loudness which create the forms of intonation. The overall loudness of an utterance is calculated by determining the intensity of all its syllables and determining their mean value, indicated in decibels (dB).

4.1.3 Overall tempo

Overall tempo is defined as the mean tempo of a given stretch of speech, excluding the brief variations in speech tempo which signal the forms of intonation. The overall tempo of an utterance may be described in terms of two phenomena, *speech rate* and *rate of articulation*. *Speech rate* is calculated by dividing the duration of the utterance, including silent periods (*pauses*), by the number of words, syllables, or segments, contained therein. In the determination of *articulation rate*, which follows the same general principle, pauses are excluded from the calculation by subtracting their total duration from the duration of the utterance⁴. While *speech* and *articulation rate* are commonly indicated as the number of syllables per unit of time, in the experimental part of this work, they will be indicated in terms of syllable/segment duration, in milliseconds (ms.)⁵.

¹ The paratone is defined as the intonational equivalent of the paragraph. A distinction is made between major and minor paratones, the latter a sub-unit of the former. (Crystal 1969, Lehiste 1975, Yule 1980, Couper-Kuhlen 1983, 1986)

² The following discussion will thus be largely restricted to the English language. While Scherer's research results often include findings for German, the only literature available on Chinese so far are the short passages provided in Chao (1968: 41), to be included here.

³ The related feature of *overall pitch-range* is defined as the overall, or mean, pitch-range covered in a given stretch of speech, indicated in semitones or Hz. As such, it corresponds to the pitch/frequency difference between the *top* and *bottom line* of utterances. Functioning as an indicator of emotional involvement and intensity, overall pitch-range also depends to a certain degree on overall pitch: All other things being equal, the overall pitch-range of utterances with a high overall pitch is narrower in absolute terms (in Hz) than in utterances with a low overall pitch.

⁴ An alternative method of calculation is to determine the number of syllables or segments per unit of time, usually one second.

⁵ See Chapters 8 and 9 below.

4.1.4 Fluency

Fluency is definable in broad terms as the term commonly used to describe the smooth and steady flow of speech marked by the absence of speech discontinuities such as pauses, hesitations⁶, repetitions and speech errors⁷, whose presence gives rise to *non-fluency*.

Once the terms *fluency* and *non-fluency* are applied to utterances, however, a more precise definition is required, taking into account the phonological status of pause phenomena, i.e. pauses and syllable-lengthenings. In this work, an utterance is regarded as *fluent* if it contains all but only those pauses/lengthenings required for the signaling of - minor and major - intonation-group boundaries, these pause phenomena generally referred to as *phonological*. Conversely, an utterance is referred to as *non-fluent* if it is marked either by the absence of one or more phonological pauses/lengthenings or by the presence of non-phonological pauses/lengthenings, i.e. pause phenomena other than those involved in phonological processes.

Two different kinds of *non-fluency* are differentiated here, based on the auditory impression, or effect, they give rise to: While utterances marked by the absence of phonological pauses and lengthenings, giving rise to an overly fast and hurried auditory impression, are referred to here as *hyperfluent*, utterances containing non-phonological pause phenomena and other speech discontinuities which slow down the speech rate and give them a hesitating or disjointed feeling are termed *disfluent*⁸. By the same definition, phonological pause phenomena of excessively long duration are also referred to here as features of *disfluency*, while overly short phonological pause phenomena not compensated for, for example by prepausal syllable-lengthening, count as features of *hyperfluency*. Finally, both insufficient and absent preboundary syllable-lengthening are referred to here as features of *hyperfluency*, and overly strong utterance-final slow-down is treated as a feature of *disfluency*.

Based on these definitions, the degree of *dis-* or *hyperfluency* of a given utterance is defined here in terms of the number and duration of speech discontinuities and/or non-phonological pauses/lengthenings, and the number of missing or overly short phonological pauses/lengthenings, respectively⁹. The above discussed vocal phenomena which give rise to *dis-* and *hyperfluency* are summarized in Fig. 37¹⁰:

<i>Disfluency</i>	<p>Overly long phonological pause/lengthening</p> <p>Non-phonological pauses/lengthenings</p> <p>Other speech discontinuities (repetitions, errors)</p> <p>Overly strong utterance-final slow-down</p>
<i>Hyperfluency</i>	<p>Overly short phonological pause/lengthening</p> <p>Missing phonological pause/lengthening (not compensated for)</p> <p>Missing utterance-final slow-down</p> <p>Insufficient utterance-final slow-down</p>

Fig. 37 The vocal phenomena which give rise to disfluency and hyperfluency

⁶ Hesitations are defined as pauses preceded by syllable-lengthening (or preboundary slow-down), while pauses not preceded by syllable-lengthening are referred to as (self-) interruptions.

⁷ Speech errors are phenomena such as false starts, mispronunciations and slips of the tongue etc.

⁸ However, *hyperfluency* is not present if the absence of a phonological pause is compensated for by preboundary syllable-lengthening, giving rise to a normal, unmarked, auditory impression.

⁹ This method of description is based on the auditory impression created by *hyper-* and *disfluency* features. The auditory impression of *dis-* and *hyperfluency* becomes stronger as the number and/or duration of these vocal phenomena increases.

¹⁰ Degree of *dis-* and *hyperfluency* is not included in Fig. 37.

4.2 *The communicative functions of overall pitch, loudness, tempo and fluency*

Overall pitch, loudness, tempo and fluency can contribute, both singly and jointly, to the communication of linguistic, extralinguistic and paralinguistic information. While in the following discussion, the para- and extralinguistic functions will be limited to English and German, with regards to their linguistic functions, a number of points relating to Mandarin Chinese are to be found in the literature and will be included here.

4.2.1 *The linguistic functions - discourse cohesion*

Similar to the forms of intonation, in their linguistic function overall loudness, pitch, tempo and fluency in English, German, and judging from the literature also in Chinese, serve in the signaling of discourse cohesion. However, while the forms of intonation signal a more *local* type of cohesion - within individual utterances and across utterance-boundaries, the overall features signal cohesion on a higher level, communicating the relative importance, or relevance, of individual utterances with respect to the speaker's message as a whole. Thus, for example, English and German utterances containing information which the speaker sees as important and wishes to emphasize are commonly marked by raised loudness and/or elevated pitch, while parenthetical remarks, containing subsidiary, qualifying, information, are often spoken with reduced loudness and pitch, optionally accompanied by a speech tempo slightly faster than in the main utterance (Crystal 1969:152 f.)¹¹. While largely the same linguistic functions apply to the overall features in Chinese, it seems that the details of their use differ slightly here and there. Thus, while parenthetical remarks in Chinese are also marked by reduced pitch/loudness and in many cases an increase in tempo (Chao 1968:43)¹², informal observations by this writer suggest that reduced speech tempo may also be encountered here. Another difference noted in the literature is the fact that elevated loudness is not normally a marker of emphatic speech in Chinese (Bolinger 1989). According to Bolinger, this feature is generally interpreted by native speakers as a sign of anger or loss of control (1989: 61)¹³. This is confirmed by Chao (1968), who notes that emphasis in Mandarin Chinese is based in the main on increases in pitch-range and duration, elevated loudness being used to communicate anger or impatience (1968: 44).

4.2.2 *The paralinguistic functions*

Regarding the paralinguistic functions of the overall features, research in social psychology and vocal emotion research, such as Allport & Cantril 1934, Fairbanks & Pronovost 1939, Kramer 1963, Williams & Stevens 1972 and Scherer et al. 1973, have established the important role played by overall pitch, loudness, tempo and fluency in the communication of emotional information, via the action of the autonomous nervous system. According to Scherer (1979a), the arousal of the sympathetic nervous system which characterizes emotional states has two main effects on vocalization and speech: (1) an increase in overall muscle tonus, leading to a rise in fundamental frequency and (2) changes, and sometimes disruptions, in the co-ordination and rhythms of reciprocal inhibition and excitation of muscle systems. In conclusion, 'It is to be expected that disruption of neuro-muscular coordination [as encountered in emotional states] affects pitch and loudness variability, stress patterns and intonation contours, speech rate (in terms of the length of phonemes as well as the number and length of pauses), the onset and decay of phonations, the precision of articulation, rhythms of speech, and many other speech parameters' (1979a:501f). On the basis of these early discoveries and observations, a large number of follow-up studies, such as Scherer 1981, Scherer et al. 1984,

¹¹ For the signaling of parentheses in German, see Sievers 1893, Riesel 1963 and Isacenko & Schädlich 1970.

¹² As another difference between Chinese and Western languages, the observations of this writer suggest that both elevated and reduced pitch-height can serve as a marker of parenthetical remarks in Mandarin Chinese.

¹³ Bolinger quotes Edwin T. Hall (1955) who observes that in contrast to Americans, Chinese do not raise their voices when emphasizing a point (1989: 61).

Ladd et al. 1985, Scherer 1986, 1995, Banse & Scherer 1996 and Leinonen et al. 1997, have examined the manner in which emotions are signaled by the overall features. Their findings are briefly outlined in the following¹⁴.

According to Banse & Scherer (1996), the methodological procedure followed in most cases has been to determine the accuracy with which emotions are inferred from speech samples¹⁵ and to examine the vocal expression patterns for specific emotions (1996: 614). With respect to the first point, results have been encouraging, since the accuracy with which listeners have been found to recognize different emotional states generally lies between 50 and 60%¹⁶. With respect to the second point, however, progress has been less satisfactory. Though the contribution of pitch, loudness and tempo in the communication of emotion is not in doubt, research results are often contradictory with respect to the precise role of these features in the signaling of specific emotional states. To name a few examples, according to the review provided by Murray & Arnott 1993, there seems to be no agreement as to whether anger is marked by increased pitch (Fairbanks & Pronovost 1939, Williams & Stevens 1972, Frick 1986) or mid/normal pitch (Fónagy & Magdics 1963, Öster & Risberg 1986). And while Fónagy & Magdics 1963 and Davitz 1964b describe joy and happiness in terms of elevated tempo, this is contradicted by Tartter (1980) and Öster & Risberg (1986) who observe reduced tempo for happiness¹⁷. Results do, however, show a strong agreement with respect to the signaling of degree of emotional arousal, corresponding to the intensity of emotions. While conflicting results are sometimes reported for pitch-range (Protopapas & Lieberman 1997) and tempo (Banse & Scherer 1996), the role of pitch, pitch-variability and especially loudness in the signaling of emotional arousal seems firmly established, high values for all three features correlating with high arousal, strong emotions and strongly developed attitudes (Scherer 1995: 239). In contrast to the intensity of emotions, however, it appears that neither their quality nor their valence, corresponding to their positive or negative feeling or undertone¹⁸, is reliably communicated by the overall features. Scherer's hypothesis (1986, 1995) that the signaling of the valence and quality of emotions depends to a strong degree on voice quality has found confirmation in a number of more recent studies, such as Klasmeyer & Sendlmeier 1995, Banse & Scherer 1996, Leinonen et al. 1997 and Johnstone & Scherer 1999.

4.2.3 The extralinguistic functions

Research in social psychology has found that overall pitch, loudness, tempo and fluency, the presence of non-phonological, hesitation pauses, play an important role in judgements of speaker personality.

4.2.3.1 Pitch and pitch-variability

With respect to pitch, the results of early studies are not always in agreement. Scherer 1970, 1974, Scherer et al. 1973, Ekman et al. 1976 and Kühnen 1977 found elevated overall pitch to be associated with dominance and assertiveness in male American speakers and with discipline and dependability in female American and male German speakers (Scherer 1979b: 155 f.). On the other hand, the results of Aronovitch 1976, Brown et al. 1975 and Apple et al. 1977, confirming earlier findings by Fairbanks & Pronovost 1939, Williams & Stevens 1972

¹⁴ Detailed reviews of the literature are also found in Murray & Arnott 1993, van Bezoooyen 1984 and Davitz 1964c.

¹⁵ In general these speech samples, which take the form of letters, numbers, nonsense syllables or standardized sentences, are produced by actors, amateur or professional, who are asked to portray different emotions while speaking.

¹⁶ Scherer (1995) does add, however, that only few emotions have been examined and even among these, not all are identified equally well: Sadness and anger are recognized best, fear and joy less well, and recognition in the case of disgust is barely above chance (1995: 237).

¹⁷ Leinonen et al 1997 also found elevated pitch in anger. While this is confirmed (for irritation) in principle by the results of Johnstone & Scherer 1999) an even higher pitch was found here for happiness.

¹⁸ The valence of an attitude or emotion is conceivable as its warm/positive or cool/cold/negative underlying feeling, or undertone. Examples of positive emotions are joy, admiration and positive surprise, emotions with a negative undertone/valency are anger, scorn and disgust. In the semantic differential modal of Osgood et al 1957, valence roughly corresponds to *evaluation*, defined as the pleasantness or unpleasantness of personal relations (Couper-Kuhlen 1986: 176).

and Uldall 1964, indicate a correlation between elevated pitch in male and female speech and submissiveness, immaturity and even emotional instability. While, as Scherer suggests, the incompatibility of these results may be due to the fact that pitch rises due to psychological arousal/stress and discipline/assertiveness are accompanied by different additional acoustic parameters which also play a role in personality judgements (1979b:187), another factor which may conceivably contribute to these contradictory results is the degree of elevation with respect to the speaker's habitual pitch¹⁹. This is suggested by the results of Terango 1966, Brown et al. 1975 and Apple et al. 1977, which indicate that negative judgements of male speakers (including weakness, effeminacy and incompetence) are particularly frequent in cases when their pitch approaches the female range, an elevation of pitch which is almost certainly greater than the pitch fluctuations encountered in everyday speech.

Finally, further support for an association between elevated pitch and *submissiveness*, is found by Ohala (1983, 1984, 1994). Based on the findings of Bolinger (1964, 1978), Ultan (1969), Cruttenden (1981) and the results of his own research, Ohala comments on the widespread, possibly universal, tendency in human communication to associate high or rising pitch with 'deference, politeness and submission', while low or falling pitch is used to convey 'assertiveness, aggression, confidence and threat' (1994: 327)²⁰. In his view, this use of pitch is part of a 'frequency code' found not only in human communication but also in the vocal signals of other species (Darwin 1872, 1874, Ekman et al. 1969, Morton 1977), and employed to convey an impression of the physical size (and strength) of the speaker. Given the generally observed link between lower pitch (as well as lower formant frequencies) and greater physical size and strength on the one hand, higher pitch (and raised formants) and smaller physical size on the other, Ohala suggests that low pitch may be adopted purposely in battle or mating to signal aggression, dominance and threat, while high pitch, associable with smaller physical stature and strength, is used to signal non-threat and submissiveness and gain the empathy and goodwill of the listener (Ohala 1994: 340 f.)

Turning now to pitch variability, results here appear a little more in agreement, most early studies, such as Addington 1968, Aronovitch 1976 and Brown et al. 1975, as well as Scherer 1979b, reporting a link between pitch variability and attributions of dynamism, extrovertedness and benevolence²¹. However, exceptions to this pattern are also encountered and this feature also seems to be associated with negative personality traits: Terango 1966 and Bennet & Weinberg 1979 report that higher pitch-variability and wider pitch-range are seen as typical features of female and effeminate male speech. According to McConnell-Ginet (1983), the negative connotation of pitch-variability in male speech may be due to the generally observed association of this feature in female speech with emotionality, mental instability and unpredictability (1983: 77f.).

4.2.3.2 Overall loudness

According to Scherer's (1979b) review of the literature, a correlation between vocal intensity and extrovertedness, or extroversion was first suggested by Allport & Cantril 1934 and confirmed by Mallory & Miller 1958 and Trimboli 1973. The results of Scherer's (1974) own investigation of this feature also support these findings and point to a correlation between vocal effort and extroversion, sociability and emotional stability (1979b: 158). Scherer's results do not, however, confirm an earlier association between degrees of loudness and the

¹⁹ In the attempt to explain these results, Scherer suggests a link between these four personality traits and a habitually high degree of arousal, this constant 'state of readiness' leading to elevated levels of pitch (1979: 155). While in assertiveness and dominance, the individual's energy is directed outwards and serves in achievement and self-enhancement, in discipline and dependability it is directed inwards, serving to maintain discipline, self-control and adherence to rules.

²⁰ Ohala's research results thus agree with those of Brown et al. 1974 and Apple et al. 1979. In his opinion, the discrepancies between their results and Scherer's (1973) are only superficial and originate in differences in experimental and measurement procedures (1994: 327).

²¹ According to Scherer (1979), however, these studies often differ in their definitions and measurements of pitch variability, so that it cannot be said for sure that their results are completely in agreement.

traits of dominance/submissiveness. In his view, this suggests that the link between extroversion and pausing behavior is not only regulated by biophysical factors, as suggested by Siegman (1978), but may also be determined by culturally-patterned norms of behavior (1979b:164).

4.2.3.3 *Speech tempo and fluency (pause phenomena)*

With respect to speech tempo and pause phenomena, research has generally focused on their potential link to extroversion and anxiety. Again, results have not always been entirely in agreement. With respect to extroversion, for example, the early association of this trait with elevated speech tempo and the relative absence of hesitation pauses (Siegman 1978) is only partly confirmed by Scherer (1970, 1974), whose results indicate fewer hesitation pauses in the speech of American extroverts but fail to find significant differences between the speech tempo of extrovert and introvert speakers (1979b:163). Interestingly, Scherer's results for German speakers point in the opposite direction: Here, a higher number of pauses was found in the speech of extroverts²². While Scherer also found no evidence for a correlation between extroversion and the duration of silent pauses, as suggested by the results of Siegman & Pope 1965, in his study, silent pauses were not actually differentiated according to length, so that a correlation between extroversion and the number of longer silent pauses still remains a possibility (1979b: 164).

Turning now to anxiety, the results of early studies, such as Murray 1971 and Rochester 1973, indicate a strong link between anxiety, slow tempo and the relative absence of silence. Scherer's findings (1970, 1974) seem to confirm this: The speech of his anxious, neurotic and generally unstable American speakers was found to be marked by a faster speech rate and fewer silent (and filled) pauses than the speech of emotionally stable subjects. On the other hand, Scherer's results are not in agreement with Rochester 1973, whose results point to fewer unfilled as well as fewer filled pauses in the speech of highly anxious speakers.

²² In particular, Siegman & Pope's (1965) results suggest the presence of fewer *long* silent pauses in the speech of extrovert American speakers.

5. Forms and functions in voice quality communication

The following discussion of the forms and functions of voice quality will be based on the articulatory *setting*¹, generally regarded as the *communicative form* of voice quality communication and defined as a configuration of muscular adjustments created in the laryngeal or supralaryngeal part of the vocal apparatus. After a brief presentation of the most important articulatory *settings* defined by Laver (1980), the communicative functions of these *settings* will be discussed².

5.1 The communicative forms of voice quality - the articulatory settings

Laver's (1980) model of voice quality differentiates between three major types of *setting*: (1) *supralaryngeal settings*, created by modifications in the supralaryngeal vocal tract; (2) *phonatory settings*, corresponding to different phonation modes of the vocal folds³ in the larynx⁴; and (3) *tension settings*, involving different degrees of tension throughout the vocal apparatus. Each of these types contains a varying number of individual *settings*, describable in terms of three criteria: (1) articulatory, relating to the muscle systems participating in their creation, (2) acoustic, pertaining to their acoustic correlates, and (3) auditory, relating to their auditory effects. In the following description of the articulatory settings, the focus will be on their acoustic correlates. For the *supralaryngeal settings*, these are the formant frequencies, for the *phonatory settings* the acoustic correlates relate to their characteristic pitch-range, closed quotient (CQ)⁵, the shape of the *glottal wave-form*, and the slope of the glottal spectrum (*glottal slope*). We begin with the basic configuration, referred to as the *neutral setting*, by reference to which all other *settings* are described.

5.1.1 The neutral setting⁶

The most important articulatory correlates of Laver's *neutral setting*, whose mode of vibration is referred to as *modal voice*, are as follows: (1) the lips are not protruded, (2) the larynx is neither raised nor lowered, (3) the supralaryngeal vocal tract is nearly in equal cross-section along its full length and there are no constrictions, (4) the jaws are neither closed nor unduly open, and (5) the vibration of the vocal folds is regularly periodic, with moderate longitudinal and adductive tension and medial compression⁷ (1980: 14 f.). The auditory correlates are the following: (1) there is audible nasality only when necessary for phonological purposes, (2) the vibration of the vocal folds causes no audible friction, and (3) the sound which is emitted

¹ The voice quality *setting* is written in cursive script to distinguish it from the ethnographic concept of the setting (See Chapters 1, 6 - 9).

² Strictly speaking, given the interconnectedness of the various muscular systems involved in the production of speech, individual *settings* rarely function independently of each other but interact in various ways, a phenomenon referred to as *acoustic interdependence*. (Laver 1980: 18 f.). While so-called compatible *settings* enter enabling relationships, incompatible *settings* inhibit each other. To name two examples, the opening of the jaws, as in the mandibular *setting*, has an enabling effect on the labial *setting* of retroflexion, whereas the phonatory settings of *modal voice* and *falsest*, which make use of different areas of the glottis, are incompatible and therefore do not occur simultaneously. In Laver's model of voice quality mutually compatible phonatory *settings* which occur simultaneously form *compound settings*. A few examples of such compound *settings* will be included below.

³ The vocal folds, also known as the vocal cords or bands, are two muscular folds situated in the larynx. Their periodic vibration, referred to as phonation, produces *voice*, the most important sound source of speech. While the frequency rate of vocal fold vibration determines pitch, the manner in which the vocal folds vibrate corresponds to the different phonatory *settings* of voice quality. The space between the vocal folds is referred to as the *glottis*.

⁴ The larynx is a tube-like structure located at the top end of the trachea, consisting of a group of cartilages, most important among which are the thyroid, the cricoid and the paired arytenoid cartilages. These are connected to each other and the rest of the vocal apparatus by the intrinsic and extrinsic laryngeal muscles, which enable them to move relative to each other. These movements give rise to the various laryngeal *settings* and create different kinds/degrees of tension within the larynx, both factors which determine the manner in which the vocal folds vibrate in phonation. See Fig. 39 below.

⁵ The closed quotient (CQ) corresponds to the relative length of the closing phase in relation to one whole vibratory cycle, consisting of (1) the opening phase in which the vocal folds pull apart increasing the glottal opening, (2) the closing phase during which the folds come together, and (3) the closed phase, during the glottis is closed. See Fig. 40 below.

⁶ Laver's definition of the *neutral setting* is 'a constellation of co-occurring *settings* in different parts of the vocal apparatus, each of which constitutes the neutral reference for the description of other[non-neutral] *settings* at that location' (1980: 14).

⁷ Adductive and longitudinal tension and medial compression are the three most important tensions created by the action of the intrinsic laryngeal muscles. More will be said of these below. See also Figs. 38 and 39.

under these conditions corresponds to [ə]. Finally, the acoustic specifications of the neutral *setting* are: (1) the first three formant frequencies of the articulated sound [ə] are located at 500, 1500 and 2500 Hz for the average male speaker, (2) the pitch-range for the average male speaker lies between 60-240 Hz, (3) laryngeal excitation is regular in frequency and amplitude, (4) the glottal waveform is roughly triangular in shape, (5) the *closed quotient (CQ)* is approximately 33%, and (6) the spectral slope of the glottal waveform descends at a rate of -10dB to -12dB per octave.

5.1.2 The supralaryngeal settings

The supralaryngeal *settings* are modifications of the neutral configuration which take place in the supralaryngeal part of the vocal apparatus, i.e. in the oral, naso-pharyngeal, pharyngeal and laryngo-pharyngeal cavities (See Fig. 38). Laver (1980: 23 ff.) distinguishes between three subtypes: (1) *longitudinal settings*, resulting from modifications of the longitudinal axis of the vocal tract, (2) *latitudinal settings*, involving modifications of its latitudinal, cross-sectional axis, and (3) *velopharyngeal settings*, involving the activity of the velum.

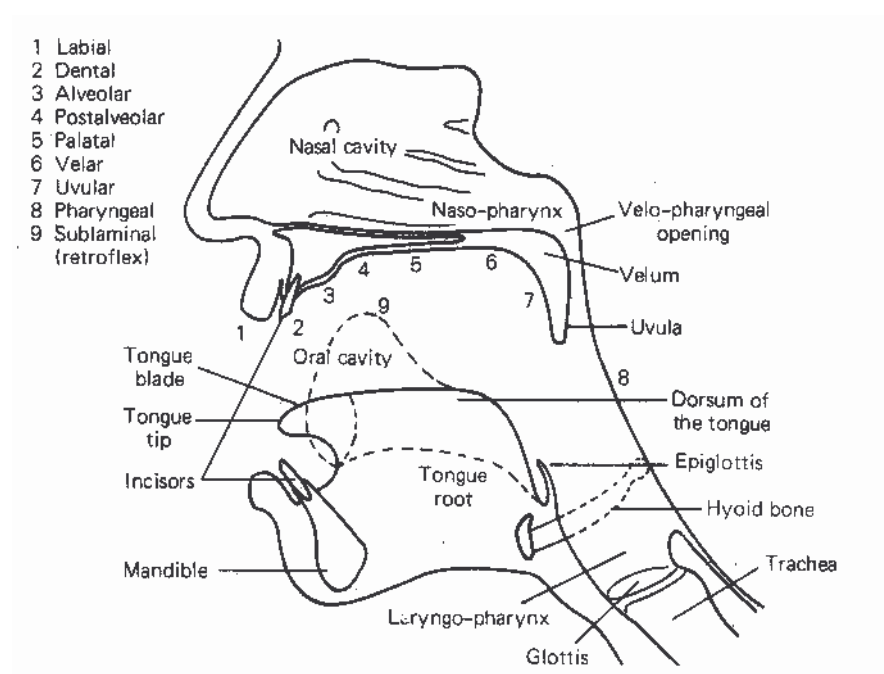


Fig. 38 The major areas of articulation (Adopted from Clark & Yallop 1990: 47)

5.1.2.1 The longitudinal settings⁸

Modifications of the longitudinal axis of the supralaryngeal vocal tract may be produced by the following four articulatory processes: (1) the raising or (2) lowering of the larynx, giving rise to the *settings* of *raised* and *lowered larynx voice*, (3) the protrusion of the lips, producing the setting referred to as *labial protrusion* and (4) the raising and retraction of the lower lip, creating *labiodentalized voice*. The vertical movements of the larynx which give rise to *raised* and *lowered larynx voice* are caused by up- and down movements of the *hyoid bone*, from which the *larynx* is suspended. These incur de- and increases in the length and width of the *laryngo-pharynx*, which in turn lead to the raising and lowering of formant frequencies⁹. According to Laver, in *raised larynx voice*, formant frequencies are raised, while in *lowered*

⁸ The longitudinal *settings* are discussed in Laver 1980: 22 ff.

⁹ Broadly speaking, a lengthening of the vocal tract, as caused by labial protrusion or the lowering of the larynx, results in a lowering of formant frequencies, while a shortening of the vocal tract, created by the retraction of the lips or the raising of the larynx, gives rise to a raising of formant frequencies.

larynx voice there is a lowering of formant frequencies. The acoustic correlates of *lip protrusion*, which involves the activity of the muscles around the mouth, and *labio-dentalized voice*, brought about by the retraction and raising of the lower lip, is a raising and lowering of all formant frequencies, respectively.

5.1.2.2 The latitudinal settings¹⁰

The *latitudinal settings* are created by constrictions and expansions of the cross-sectional area of the supralaryngeal vocal tract. Laver distinguishes 5 types of *latitudinal setting*, the *labial*, *lingual*, *faucal*, *pharyngeal* and *mandibular settings*, named after the articulator principally involved in each case.

The 8 *labial settings* are defined in terms of horizontal or vertical constriction/expansion of the interlabial space (Laver 1980: 36). Constriction of interlabial space leads to a lowering of formant frequencies, while expansion incurs a general raise of formant frequencies. The eight *lingual settings* are: *dentalization*, *alveolarization*, *palato-alveolarization*, *palatalization*, *velarization*, *uvularization*, *pharyngealization* and *laryngo-pharyngealization*. In the first 4 *settings*, sometimes also referred to collectively as ‘fronted speech’, the tip and blade of the tongue are brought up- and forwards towards the upper incisors, the alveolar ridge or the hard palate, with the acoustic effect of raising all formant frequencies, in particular F2. The latter four *settings*, also referred to as ‘backed speech’ involve a ‘backing’ movement of the tongue, whose body is retracted towards the soft palate and the uvula, incurring a lowering of F2 and a raising of F1. Coming now to the *faucal*, *pharyngeal* and *mandibular settings*, both the *faucal* and *pharyngeal settings*, involving constrictions of the pharynx by the *faucal pillars*¹¹ and the muscles of the pharyngeal walls, respectively, have the acoustic effect of raising F1, a concomitant lowering of F2 being found in the *pharyngeal settings*. Finally, the acoustic effect of the *mandibular settings*, which involve up and down movements of the lower jaw, is a raise of F1, which increases with the degree of jaw-opening.

5.1.2.3 The velopharyngeal settings¹²

Laver’s three *velopharyngeal settings* are the *neutral velopharyngeal setting*, *nasality* and *denasality*. They are created by the activity of the muscles in and around the *velum* which open and close the *velopharyngeal port* connecting the pharyngeal and nasal cavities¹³. In the *neutral velopharyngeal setting* the velum is raised and the velopharyngeal port is loosely closed, opening for the production of phonologically nasal segments. In *nasality* the velum is lowered and the velopharyngeal port remains more or less open all the time, giving rise to the characteristic auditory effect of this *setting*, described as ‘nasal voice’ or ‘nasal twang’ (1980: 68). Finally, in *denasality*, the velopharyngeal port remains relatively tightly closed, and hardly opens at all on phonologically nasal segments¹⁴. The acoustic specifications of *nasality* are generally agreed to be: (1) an overall loss in intensity, (2) the creation of three nasal formants at (1) 200 to 300 Hz, (2) about 1000 Hz and (3) about 2000 Hz and (3) the creation of anti-resonances, or anti-formants, one at 500 to 700 Hz, two others between 900 and 1800 Hz (1980: 91). The *denasal setting* is described acoustically in terms of the minimization of the acoustic cues of *nasality* on the effected segments (1980: 92).

¹⁰ Laver 1980:34 ff.

¹¹ The *faucal pillars* are two sets of muscular arches located at the junction of the mouth and the pharynx. The auditory effect of the latitudinal approximation of the *faucal pillars* is described by Alexander Bell as having a ‘peculiar metallic ring, similar to the tone of a brass musical instrument’ (1908:19f., quoted by Laver 1980:57).

¹² The velopharyngeal *settings* are discussed in Laver 1980: 68 ff.

¹³ The articulatory processes which underlie the opening and closing of the velopharyngeal port are complex and have not been definitely clarified. The most important articulator is, however, agreed to be the *velum*, or soft palate, which is lowered and raised to open and close the passage into the nasal cavity. On the other hand, it appears that this process is accompanied by the activity of the sphincter muscles around the velopharyngeal port. (Laver 1980: 74 f.)

¹⁴ The auditory effect of *denasality* is often described in terms of ‘speaking with a cold in the head’.

5.1.3 The phonatory settings

Laver (1980: 93 ff.) distinguishes five basic *phonatory settings*, *false alto*, *whisper*, *creak*, *harshness* and *breathiness*. These are defined in terms of their characteristic modes of phonatory vibration¹⁵, contrasted against *modal voice*, the vibration mode of the *neutral setting*. The vibration modes of the 5 phonatory *settings* result in the main from the activity of the intrinsic laryngeal muscles which, through the action of *adductive* and *longitudinal tension* and *medial compression*, modify the tension in the vocal folds and the degree of opening of the glottis¹⁶.

5.1.3.1 False alto¹⁷

In *false alto*, adductive and longitudinal tensions are high and medial compression is strong, this resulting in stiff, relatively thin vocal folds, with only the vocal ligaments along their edges participating in phonatory vibration (Laver 1980: 118). The glottis remains slightly open and subglottal air-pressure is lower than in *modal voice*, so that this *setting* has a slightly ‘whispery’ auditory effect¹⁸. The most important acoustic specifications of *false alto* are: (1) a high pitch-range, between 275 and 634 Hz (for male voices), (2) a steep glottal waveform and (3) a steeply sloping glottal spectrum, which falls at a rate of about -20 dB per octave.

5.1.3.2 Whisper¹⁹

The most characteristic articulatory feature of *whisper* is agreed to be the triangular-shaped opening in the cartilaginous part of the glottis, giving rise to the characteristic ‘whispery’ quality of this tone of voice. Adductive and longitudinal tension are low, medial compression is medium to high (Laver 1980: 121). Among the acoustic correlates of *whisper* are: (1) a steep, ‘spiky’, glottal waveform and (2) a relatively steep spectral slope, a feature this *setting* shares with *false alto*. When *whisper* combines with *modal voice*, this gives the compound phonation mode of *whispery voice*.

5.1.3.3 Creak²⁰

According to Laver, the articulatory correlates of *creak* are slack, but thick vocal folds which, combined with the strong damping effect of the adducted *ventricular folds*²¹, give rise to the characteristic auditory effect of this *setting*, described by Catford in terms of ‘a rapid series of taps, like a stick being run along a railing’ (1964: 32)²². There is low longitudinal tension, strong adductive tension and strong medial compression, and subglottal air pressure appears to be lower than in modal voice (Laver 1980: 124). According to Laver, the most important acoustic correlates of *creak* are: (1) a relatively low pitch-range, between 20 and 90 Hz for the average male speaker, (2) a high degree of pitch variation (*jitter*) and (3) a highly irregular glottal waveform. (See Fig. 39)

¹⁵ Phonation, i.e. vocal fold vibration, results from a combination of aerodynamic and physiological forces. The most widely accepted theory of vocal fold vibration is the *aerodynamic myoelastic theory* (van den Berg 1958, 1968), which incorporates the aerodynamic forces, muscle activity and vocal fold tissue structure and elasticity (Broad 1979, Hirano & Kakita 1985). A detailed account of the aerodynamic myoelastic theory is found in Laver 1980: 95 f. and Clark & Yallop 1990: 36 f.

¹⁶ The three most important groups of intrinsic laryngeal muscles are (1) the cricothyroid muscles, the thyroarytenoid muscles and the vocalis muscles, which regulate vocal fold tension (longitudinal tension), (2) the posterior cricoarytenoid muscles, which control the opening (abduction) of the glottis, and (3) the lateral cricoarytenoid muscles and the interarytenoid muscles, which control its closure (adductive tension). Medial compression, which closes the ligamental, but not the cartilaginous part of the glottis, is created by the action of lateral cricoarytenoid muscles and reinforced by the lateral parts of the thyroarytenoid muscles (Laver 1980: 108).

¹⁷ False alto is discussed in Laver 1980: 118 f.

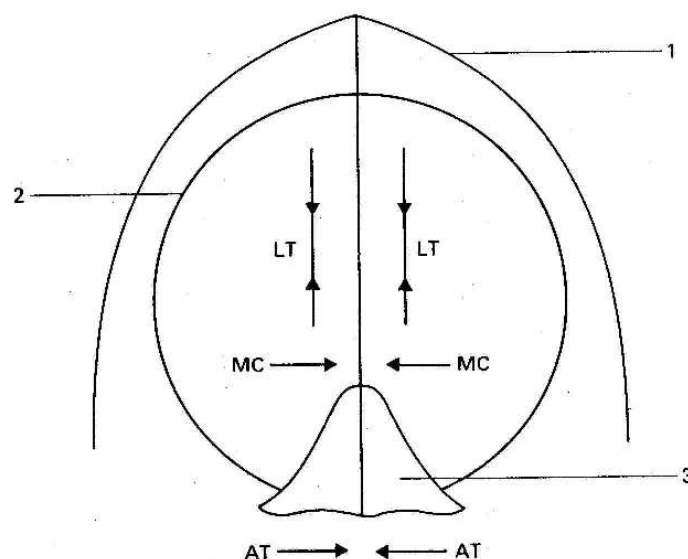
¹⁸ With respect to the auditory effect of *false alto*, Zemlin (1964: 155) is quoted as noting that the quality of tone produced by *false alto* is ‘almost flute-like in nature’ (Laver 1980: 120).

¹⁹ Whisper is discussed in Laver 1980: 120 ff.

²⁰ Creak, also referred to as *laryngealization*, *vocal* or *glottal fry* in the literature, is discussed in Laver 1980: 122 ff.

²¹ The ventricular folds, sometimes referred to as the ‘false vocal folds’, are positioned within the larynx slightly above the (true) vocal folds. The ventricular folds do not participate in normal phonation, but activity of the ventricular folds may accompany vocal fold vibration in certain settings, such as creak (adduction) and harshness (vibration).

²² Cited in Laver 1980: 124.



LT - Longitudinal tension	1. Thyroid cartilage
MC - Medial compression	2. Cricoid cartilage
AT - Adductive tension	3. Arytenoid cartilage

Fig. 39 Types of laryngeal tension²³

5.1.3.4 Harshness²⁴

The physiological correlates of *harshness* are extreme adductive tension, strong medial compression and relatively strong longitudinal tension. Its characteristic auditory effect, commonly described as ‘raucous’, ‘strident’ and ‘rough’ (Laver 1980: 127) originates in the high degree of overall laryngeal tension, leading to strong pitch variation (*jitter*) and the presence of aperiodic noise. The acoustic specifications of *harshness* are: (1) a relatively low pitch-range, around or slightly above 100 Hz in the male voice, (2) a relatively high intensity, (3) the strong presence of aperiodic noise, (4) a high degree of pitch variation, (5) an asymmetric glottal waveform with an abrupt closing section, (6) a high closed quotient (*CQ*) and (7) a gently falling, ‘flattened’ spectral slope with strong upper harmonics and falling at less than -12dB per octave.

5.1.3.5 Breathiness²⁵

The characteristic auditory effect of *breathiness* is described by Catford (1977: 99) as resembling ‘the sound of voice mixed with breath, [...] somewhat like that of sighing.’ As he adds, ‘The vocal folds are vibrating, but never closing or, indeed, coming anywhere near closing. They simply ‘flap in the breeze’ of the high velocity air-flow’ (1979: 99)²⁶. The articulatory correlates of *breathiness* are: minimal adductive tension, weak medial compression and low longitudinal tension. The glottis remains slightly open along most of its length and the vibration of the lax vocal folds is inefficient and accompanied by slight audible friction. As overall tension in the vocal folds is low and part of the acoustic energy is lost as a result of glottal opening, pitch and intensity are relatively low in this *setting*. The acoustic correlates of *breathiness* are: (1) a low closed quotient (*CQ*), (2) a rather steep glottal waveform, (3) a relatively steep glottal spectrum, falling at a rate of more than -12 dB per octave, and (4) the presence of aperiodic noise in the spectrogram. The combination of *breathiness* with *modal*

²³ Adopted from Laver 1980: 109.

²⁴ *Harshness*, also referred to as *tense* or *pressed voice* in the literature, is discussed in Laver 1980:126 ff.

²⁵ *Breathiness* is discussed in Laver 1980: 132 ff.

²⁶ Quoted in Laver 1980: 132.

voice gives rise to the compound phonation mode of *breathy voice*. Fig. 40 shows the action of the posterior and lateral cricoarytenoid muscles in the opening and closing of the *glottis*, i.e. the abduction and adduction of the vocal folds formed by the vocalis and thyroarytenoid muscles.

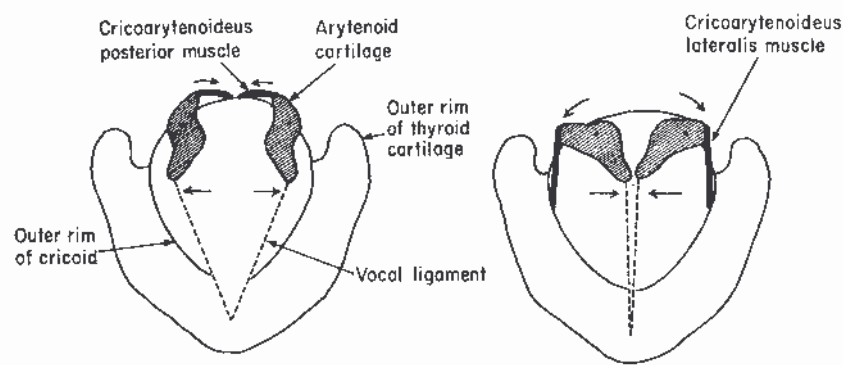


Fig. 40 View of the glottis from above²⁷

5.1.4 The tension settings²⁸

In Laver's model of voice quality, the supralaryngeal and phonatory *settings*, all specific to particular areas of the vocal apparatus, are distinguished from what he describes as *settings* of 'overall degree of muscular tension', whose effect on the vocal system is broader, including both the larynx and the supralaryngeal vocal tract (1980: 141). Laver distinguishes between *tense* and *lax voice*, characterized by a high and low degree of overall muscular tension, respectively.

5.1.4.1 Tense voice

Tense voice is characterized by the *tense, harsh* phonation mode described above for *harshness*²⁹. It is accompanied by a high degree of tenseness in the supralaryngeal part of the vocal tract, extending from the upper larynx to the walls of the pharyngeal and oral cavities and giving rise to a segmental articulation which is described as vigorous and effective (Laver 1980: 155). Among the acoustic specifications of *tense voice* are thus: (1) a relatively high intensity, (2) a high closed quotient (*CQ*), (3) an asymmetric glottal waveform with an abrupt closing section, and (4) strong upper harmonics in the speech spectrum, giving rise to a 'flattened' glottal slope, descending at less than -12 dB per octave. The auditory impression, or effect, of *tense voice* is commonly described as 'rough', 'strident' or 'hard'.

5.1.4.2 Lax voice

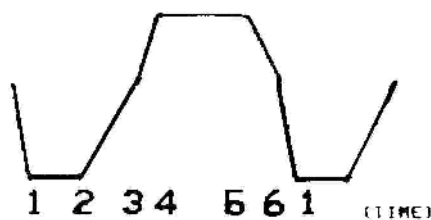
Lax voice is characterized by lax, gentle phonation, with a soft and relatively low-pitched auditory effect. Subglottal pressure and general tension are low throughout the vocal system and articulatory movements are gentle, with lesser articulatory effort than in *modal voice*. Articulatory targets are often not attained, leading to slighted consonants and neutralized vowels (Laver 1980: 153). The auditory effect of *lax voice*, which is also often accompanied by slight *nasality*, is commonly described as 'soft', 'muffled' or 'dull'. Similar to what was discussed above for *breathiness*, adductive and longitudinal tension are slight and medial

²⁷ Fig. 39 is from Hardcastle 1976:77.

²⁸ The tension *settings* are described in Laver 1980: 141 ff.

²⁹ Based on the strong similarities in their auditory and acoustic correlates, the phonatory *setting* of *harshness* is here equated with the phonatory vibration mode of Laver's *tense voice*. This also appears to be the general practice in the literature.

compression is also lower than in *modal voice*. It does seem, however, that overall laryngeal tension in *lax voice* is a little higher than in *breathiness*, which may - but need not necessarily - co-occur with *lax voice* if the glottis remains open due to an even lower degree of adductive tension and medial compression (1980: 146)³⁰. The acoustic correlates of this *setting* are similar to those described above for *breathiness* and include: (1) a relatively low intensity, (2) a low pitch-range, (3) a low closed quotient (*CQ*), (4) a steep, and more symmetrical glottal waveform than that of *tense voice*, and (5) a relatively steep glottal spectrum. (6) As there is less widening of the glottis in *lax voice* than in *breathiness*, however, there is no presence of aperiodic noise.



- 1 – 2 Complete closure of the vocal folds
- 2 – 4 The vocal folds part; Opening phase of the cycle
- 4 – 5 The vocal folds are apart. There is no lateral contact.
- 5 – 6 The vocal folds close.
- 6 – 1 Rapid increase in closure (vocal fold contact)

Fig. 41 The glottal cycle for modal voice³¹

³⁰ This is also the position of Ladefoget (1971) and Catford (1977). With respect to degree of glottal constriction, Ladefoget sees *lax voice* as an intermediate stage between *modal voice* and *breathiness* (1971: 18, quoted by Laver 1980: 147).

³¹ Fig. 41 is from Childers et al. 1984:137

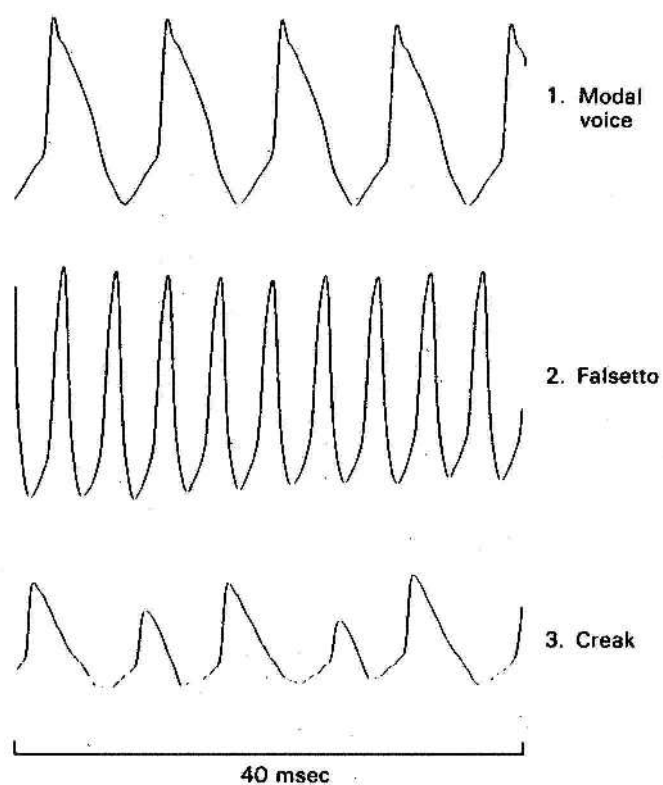


Fig. 42 Glottal waveforms of *modal voice*, *falsetto* and *creak*

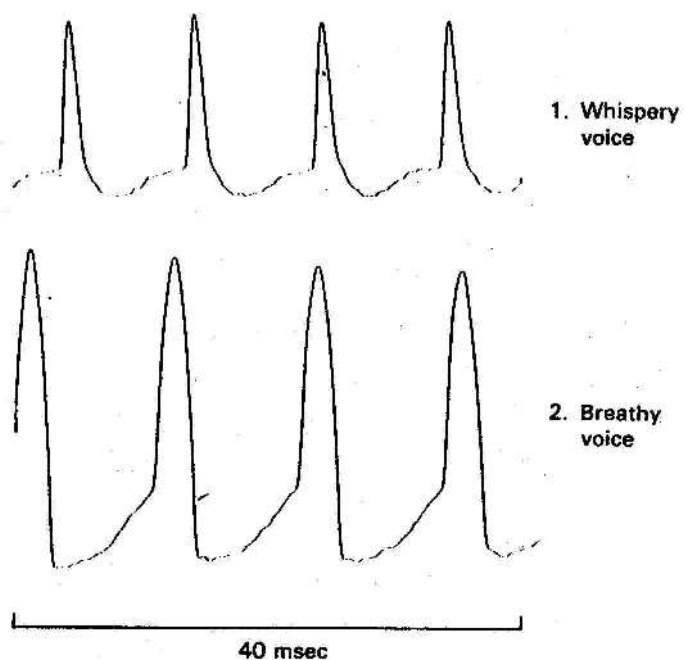


Fig. 43 Glottal waveforms of *whispery voice* and *breathy voice*³²

³² Figs. 42 and 43 are from Laver 1980, page 112 and 114, respectively

5.2 The communicative functions of voice quality

While there is growing evidence that voice quality - especially the phonatory *settings* - may also serve linguistic functions of the kind discussed above for the forms of intonation and the overall features³³, the general consensus is that voice quality contributes first and foremost to the signaling of para- and extralinguistic information. This discussion of the functions of the voice quality *settings* defined above will therefore focus on these two dimensions, the phonatory and the supralaryngeal *settings* receiving separate treatment. Given the scarcity of research into other languages, the following discussion will be mostly limited to English, but suggestions and, whenever possible, references to para- and extralinguistic voice quality functions in German and Chinese will also be included. In addition to this, a number of research results of particular interest to this work relating to other languages will also be mentioned.

5.2.1 The paralinguistic functions of voice quality

It has long been known that emotions give rise to physiological processes affecting various aspects of speech (Scherer 1979a, 1986, Scherer et al. 1984). However, the precise effect of emotional states on phonation, in terms of neurophysiologically-mediated changes to vocal fold vibration, has only recently been clarified (Klasmeyer & Sendlmeier 1995, Banse & Scherer 1996). Today it is established beyond doubt that in its paralinguistic function, phonatory voice quality plays a major role in the signaling of emotions, and studies are now focusing on the link between the quality and valence³⁴ of emotions and certain acoustic parameters used to describe phonatory voice quality, i.e. the shape of the glottal waveform and the glottal spectrum (Leinonen et al. 1997, Johnstone & Scherer 1999).

With respect to the paralinguistic functions of the supralaryngeal *settings*, research findings point in a slightly different direction. While they also participate to some degree in the signaling of emotion, their primary function is the communication of interpersonal attitude.

5.2.1.1 The paralinguistic functions of the phonatory settings

In the present subsection we will discuss *whisper*, *creak*, *breathiness*, *lax voice*, *harshness* and *tense voice*, the last two being taken together³⁵. A few remarks on the paralinguistic functions of *modal voice* will also be included.

Whisper

The most well-known paralinguistic use of *whisper* is undoubtedly the signaling of secrecy and confidentiality, a function it holds in English (Laver 1980: 122) as well as German. While the same function applies in principle to this *setting* in Chinese, the observations of this writer suggest that *whispering* is employed less commonly there than in the West³⁶. Apart from this, vocal emotion research has found *whisper* to mark extreme states of emotional arousal, such as grief and sorrow (Williams & Stevens 1972), while Laver has observed the presence of this

³³ According to Laver, for example, speakers of RP often use *creak* together with low falling intonation to signal the end of their turn in conversation (1980: 126). As a second example of a possible linguistic function, studies in phonetics have shown that rises in pitch are often accompanied by a tensing of the vocal folds, suggesting that relative *tenseness* may serve as an additional marker of stressed syllables, including the *onset* and the *nucleus* (Hardcastle 1976: 75f.).

³⁴ The term 'quality' is used in vocal emotion research to refer to different types of emotion, such as anger, surprise and joy. The 'valence' of an emotional state may be glossed as its positive or negative feeling, or undertone. Thus, anger, hate and contempt have a positive valence, while joy, happiness and admiration are described in terms of positive valence.

³⁵ While *falsestto* does not appear to hold any genuine communicative functions in English, German or Chinese, in the first two languages, it is often used in mimicry (Couper-Kuhlen 1996) and the portrayal of high-pitched voices, the latter particularly frequent in children's programmes, especially cartoons. This use of *falsestto* is clearly motivated by its high pitch and soft, 'flutelike' tone of voice, to which attention was called to above. A related use of *falsestto* is found in the Chinese Peking Opera, where actors use it in the portrayal of young male, as well as female, characters.

³⁶ This should not be taken to mean that whispering is not possible in Chinese, which is not the case. (See Meyer-Eppler 1957/1972)

setting in what he refers to as the ‘little girl’s voice’, in which women adopt the speech behavior and tone of voice of little girls to various communicative aims (Laver 1980: 47).

Creak

In vocal emotion research *creak* is most immediately associated with the communication of sarcasm/irony (Fónagy & Magdics 1963, Fónagy 1981) and scorn (Fairbanks & Pronovost 1939, Leinonen et al. 1997) in English, a function which it also holds in German. According to Laver, another paralinguistic function of *creak* in English conversation is the communication of bored resignation (1980: 126). Neither function seems to apply to this *setting* in Mandarin Chinese, however, where it is most commonly encountered as a purely phonetic feature of the low rising and sometimes also the high falling lexical tone (Laver 1980: 126).

Breathiness

Breathiness is often observed as a feature of positive emotions, such as joy/happiness (Fónagy & Magdics 1963) and positive surprise/admiration (Fónagy & Magdics 1963, Leinonen et al. 1997)³⁷. Though only established for the English language, these functions are also likely to apply to this *setting* in German, and possibly also in Chinese. Brown & Levinson (1987) observe that *breathiness* is commonly used by Tzeltal women to signal politeness in interactions with men. According to the observations of this writer, this also applies to the speech of Chinese (as well as Japanese) women, whose voices take on a distinctly breathy tone of voice in conversations with men, and possibly also women, on whom they wish to make a good impression. The related *setting* of *breathy voice* has been found to signal intimacy and sensuality, a function it holds in both English and German, as well as, so it seems, in Chinese³⁸.

Harshness/tense voice

The tense phonation mode of *harshness*, generally associated with a ‘harsh’, ‘hard’, ‘blaring’ and ‘strident’ auditory effect, is observed most often in anger and rage (Trojan 1952, Davitz 1964b, Fónagy 1981, Leinonen et al. 1997). Given the observation that vowels and consonants produced in a situation of fear are often more precisely articulated than in neutral speech (Williams & Stevens’ 1972: 1249), it seems possible to conclude that *harshness* may also function as a marker of fear in English. A ‘moderately blaring’ tone of voice is also reported for happiness and joy (Davitz 1964b). Finally, according to Fónagy & Magdics 1963 and Fónagy 1981, *tenseness* may also serve as a marker of irony³⁹. While the latter function does not seem to apply to German or Chinese, *harshness* in these languages is certainly associated with strong, intense emotional states, in particular those with negative valence, such as anger and rage⁴⁰.

³⁷ As two apparent exceptions to this pattern, Fónagy and Magdics 1963 and Trojan 1952 observe what they call ‘breathy’ voice quality in states of anger. However, the same emotion is also described in terms of ‘tenseness’. It therefore seems more likely that they are in fact talking about the irregular aperiodic noise (*jitter*), described above as a feature of *harshness* and originating in pitch irregularities.

³⁸ With respect to the function of *breathiness* in women’s speech, it should perhaps be added that Henton and Bladon’s 1985 association between sexual arousal and *breathiness*, which in their view is adopted consciously by women to arouse interest in male listeners, is not supported here. First of all, it is not *breathiness* but *whispery* and *breathy voice* which is associated with the signaling of ‘sensuality’ in the literature (Laver 1975:304, 1980: 298 and 235). A second factor which speaks against what would amount to an ‘interaction strategy’ of ‘feigning sexual arousal’ is the social stigma attached to such behavior. A woman striving to sound sexually aroused in everyday conversation would be more likely to attract suspicion and aversion than interest. It would, therefore, seem more appropriate to associate this use of *breathiness* with the signaling of admiration, one of the paralinguistic functions discussed for this *setting* above.

³⁹ Note, however, that it cannot be ascertained whether Fónagy (1981) is indeed speaking about tenseness. While, according to his account, particularly the ‘second phase’ of the expression of irony is marked by ‘head register’ - which according to Laver corresponds to *harshness/tense voice* - and a ‘tense articulation’, resulting in a ‘clearer and sharper vowel quality’ (1981: 65), features mentioned above as characteristic of *harshness*, on the other hand, Fónagy also observes a very high overall pitch - up to 250 Hz, for male speakers - a feature not normally associated with *harshness*.

⁴⁰ See Chao 1968: 44, where he refers to the ‘hammer-like’ tone of voice encountered in anger and impatience, a description highly reminiscent of the terms ‘strident’ and ‘blaring’ used in the literature to describe the auditory effect of *tenseness/harsh voice*.

Lax voice

In vocal emotion research, *lax voice* is associated with affection and tenderness. According to Fónagy & Magdics (1963: 294), tenderness is marked by ‘extremely soft’ and ‘a little nasal’ articulation, while Fónagy (1981: 55) notes ‘complete but smooth contact of the vocal folds’ and Davitz (1964b) describes an affectionate tone of voice in terms of ‘softness’, ‘low pitch’, ‘resonant timbre’ and ‘slurred’ enunciation. Impressionistic though they are, these descriptions would seem to fit very well to the tone of voice described by Laver as *breathy voice*.

Modal voice

Modal or *neutral voice*, also widely referred to as ‘chest register’ (Laver 1980: 110), is hardly discussed in the literature, one of the few exceptions being Laver & Hanson 1981. The generally held view appears to be that *modal voice* is a feature of emotionally neutral, ‘matter-of-fact’ speech, characterized by the absence of all emotional colouring, positive or negative (Laver 1980: 14, Murray & Arnott 1993: 1103)⁴¹.

5.2.1.2 *The paralinguistic functions of the supralaryngeal settings*

The three most well-known *supralaryngeal settings* are the - *longitudinal* - settings of *labiodentalized voice* (smiling), *labial protrusion* (found in pouting) and *nasality*. Employed to signal positive emotions, *labiodentalized voice* is described in the literature as a marker of joy, happiness (Tartter 1980), contentedness and admiration (Leinonen et al. 1997). By contrast, *labial protrusion*, often accompanied by frowning, is employed most often in the communication of negative emotions, such as irritation and anger. This *setting* may, however, also be used to signal emotionally neutral thoughtfulness and contemplation. Finally, *nasalization* is often associated with irony and sarcasm in the literature (Gibbs 1986, Fónagy 1981, Laver 1980). While the functions of *smiling* and *labial protrusion* apply to all three languages discussed here, the above described link between *nasalization* and irony/sarcasm does not seem to apply to Chinese⁴². Coming to the *latitudinal settings*, *palatalization* is widely discussed in the literature. One of its most frequent uses in English and German is described by Laver, who notes that this *setting* is encountered as a feature of *motherese* and the ‘little girl’s voice’ (1980: 47)⁴³. According to the observations of this writer, *palatalization* is also used in this manner in Chinese. It is also frequently employed by Chinese women engaged in polite social interaction, particularly with male interaction partners⁴⁴.

5.2.2 *The extralinguistic functions of voice quality*

In their extralinguistic functions, the supralaryngeal *settings* are most often associated with the signaling of speech community membership. While this function also applies, to a slightly

⁴¹ Note, however, that if ‘chest tone’ does indeed correspond to Laver’s *modal voice*, there are two exceptions to this pattern: Both Fónagy & Magdics (1963) and Trojan (1952) describe ‘chest tone’, i.e. modal voice, as one of the characteristic features of scorn.

⁴² This may be due to the fact that *nasality* is used for phonological purposes in most Chinese dialects, as for example Mandarin. According to the observations of this writer, the signaling of irony in Mandarin Chinese, especially in male speech, is often accompanied by a characteristic *setting* not yet described in the literature and involving out- and downward movements of the lips.

⁴³ What exactly is meant by this cannot be ascertained, as no explanation is offered. It would seem, however, that Laver is referring to the adoption of a ‘girlish’ tone of voice, which is often, though not exclusively, encountered in the speech of women and young girls.

⁴⁴ As a point of interest to this work, the paralinguistic functions of the *longitudinal* and *latitudinal settings* discussed here are related to and may even derive from the ethnologically based functions suggested for *smiling* and *lip protrusion*⁴⁴ by Ohala (1984, 1994) as part of the *frequency code*. The *frequency code* regards not only pitch-height and range but also the formant frequencies as indicative of (genuine or assumed) physical size and posture: Thus, high formant frequencies are commonly encountered in the voices of physically small speakers, while low frequencies are markers of large physical size. To Ohala therefore, *lip protrusion*, leading to a lowering of formants is a means of signaling a large, potentially threatening physical size, while *smiling*, incurring a rise of formant frequencies, communicates a small, non-threatening physical size and a submissive posture. Note that formant changes of the kind discussed here are also encountered in the *latitudinal settings*. A case in point is *palatalization* which also incurs a raise in F2 and F3 in comparison with the *neutral setting*. It would therefore seem that, just like the *longitudinal setting* of smiling, *palatalization* could be associated in general terms with the communication of smallness, non-threat and submissiveness (Ohala 1994: 341). Such an interpretation would be highly compatible with the paralinguistic functions observed for this *setting* by Laver - its use in the signaling of *motherese* and the ‘little girl’s’ voice, as well as its observed function in Chinese female speech.

lesser degree, to some of the phonatory *settings*, their most primary extralinguistic function appears to be the communication of personality traits. Recent research has shown that certain *supralaryngeal settings* can also contribute to judgements of speaker personality.

5.2.2.1 *The extralinguistic functions of phonatory voice quality*

According to Scherer's (1979b) survey of the literature, early research into the role of phonatory voice quality as a marker of personality focused widely on *breathiness/laxness* and *harshness/tense voice*. While *breathiness* was found to be indicative of introversion, anxiety and neurotic tendencies (Diehl et al. 1959, Moore 1939), harsh, 'metallic' and 'strident' (i.e. tense) voices were generally associated with emotional stability, extroversion and dominance (Mallory & Miller 1958, Moore 1939). As Laver notes:

We seem prepared, as listeners, to draw far reaching conclusions from voice quality about long-term psychological characteristics of a speaker, in assessments of personality. In Western culture, we are ready to believe, for example, that a harsh voice is correlated with more aggressive, dominant, authoritative characteristics, and a breathy voice with more self-effacing, submissive, meek personalities (1968: 49 - 50).

In his review of the literature on the personality connotations of *tense voice* and *lax voice*, Laver (1975) finds much the same patterns, leading him to suggest an intrinsic link between these voice qualities and variations in overall levels of muscular tension throughout the vocal system. This hypothesis is supported by Scherer (1979b), who observes that active, disciplined, controlled and dominant speakers tend to have rather *tense* voices, as opposed to the *laxer* voices of passive, submissive and relaxed speakers (1979b: 159).

5.2.2.2 *The extralinguistic functions of the supralaryngeal settings*

The association between the supralaryngeal *settings* and speech community membership has a very long history. According to Laver (1980), one of the earliest references relating to this function is Webster's (1789) association between *nasality* and American English. A little later, in the 18th century, this same *setting* is described as a characteristic feature of the speech of Prussian generals and priests (Luchsinger & Arnold 1965). Among the other *settings* discussed by Laver are *palatalization*, generally seen as a characteristic feature of the French language, and a *setting* he describes as '*velarized voice with denasality*', typical for certain parts of Lancashire' (1980: 48),⁴⁵.

With regards to the personality traits commonly associated with dialectal speech, studies conducted in Britain, such as Giles 1970, 1971 and Giles & Bourhis 1973⁴⁶, report that speakers of regional accents were generally judged to be more friendly, sociable and trustworthy than speakers of RP, who were seen as more intelligent and competent, but also more dominant and less trustworthy. Similar results have been found in more recent follow-up research, focusing on both male and female speakers. Thus, for example, Elyan et al. 1978 and Giles & Marsh 1979 found that male and female RP speakers were rated as more competent, egalitarian, and more masculine (egotistic and independent) as compared to speakers of Southern Welsh and Lancashire accents, who were, however, given higher rates for sociability and likeability⁴⁷.

⁴⁵ More examples are provided by Catford (1977: 192f.).

⁴⁶ See also reviews provided in Giles & Powesland 1975 and Smith 1985.

⁴⁷ As an interesting further result, these two studies also found that female speakers of RP were also judged to be more 'feminine', leading them to the conclusion that female speakers of RP display 'psychological androgyny', described as strongly masculine behavior in some respects and strongly feminine behavior in others (Elyan et al. 1978: 129). Evidence not in concordance with this line of thought has, however, since been presented by Giles et al 1980. (See discussion in Smith 1985: 87 ff.)

Summary and conclusions of Part Two

The following is a summary of the research findings discussed in Chapters 2 to 5 of relevance to the two main topics underlying the present discussion - the functional diversity, and ambiguity, of suprasegmental phenomena, and the manner in which the suprasegmental forms contribute to the signaling of (interpersonal) attitude.

The functional diversity of the suprasegmental forms

In the above discussion of the literature it was found that in English, German and to some extent also in Chinese the communicative forms of suprasegmental communication - the forms of intonation, overall loudness, pitch, tempo and fluency and voice quality *settings* - all possess a high degree of functional diversity, as they may contribute to the signaling of linguistic, paralinguistic and extralinguistic information. In their linguistic function both the forms of western intonation - *tonality*, *tonicity* and the nuclear *pitch-patterns* - as well as the *overall forms* signal information serving in the signaling of discourse cohesion. In their paralinguistic function the same forms communicate information relating to speaker attitude and emotion. And in their extralinguistic function the forms of intonation and the *overall forms* communicate information relating to speech community membership and certain traits of speaker personality. A similar, though perhaps slightly lesser degree of functional diversity was found for voice quality *settings*: While their linguistic functions appear to be limited, in their para- and extralinguistic functions both kinds of *setting* serve to signal attitudinal, emotional and indexical information of the kind discussed above for the other forms. Though research activity has been largely restricted to English and German, the results for the overall features and voice quality, whose functionings are based on universal physiological processes, are also applicable in principle to the Chinese language. Finally, while research into the forms and functions of Chinese intonation is still at a very early stage, the functions discussed in Chao 1968, relating mainly to attitudinal functions, certainly hint at a similar diversity on the part of these forms. It is to be expected that a closer look at the paralinguistic and linguistic functions of the Mandarin Chinese forms of intonation will confirm this proposition.

All these findings evidence a strong functional ambiguity of the suprasegmental phenomena, as every one of the discussed communicative forms can signal a large variety of different communicative meanings. Therefore, as Gumperz and Tannen have amply demonstrated, the correct interpretation of these forms in accordance with the speaker's intended message is often far from easy. In addition, the communicative use of the suprasegmental phenomena, in particular the forms of intonation, often differs between speech communities, so that interaction partners from different communities, possessing different socio-cultural, and therefore also different linguistic, background knowledge are particularly hard-pressed in inferences of speaker meaning and intent. In the absence of adequate background knowledge, confusions of the linguistic, paralinguistic and extralinguistic usage of these forms occur. Frequently, for example, suprasegmental phenomena originally intended to signal linguistic information are misunderstood in terms of paralinguistic or extralinguistic meaning. To name one frequent example, elevated intensity - to a certain extent accompanied by elevated pitch - which is originally meant as a marker of emphatic speech, is often misunderstood by the uninformed listener as a sign of anger or irritation, this in turn leading him to conclude that the speaker is an irascible, unpredictable and untrustworthy type of person⁴⁸.

⁴⁸ In this context, the 'uninformed' listener not only adheres to different communicative conventions as the speaker, but is also insufficiently familiar with the speaker himself, so that he has no adequate understanding of the 'baseline' of his voice, in terms of its long-term features of loudness, pitch, tempo, fluency and voice quality, and the speaker's habitual, characteristic use of intonation.

The signaling of the metamessage

In Chapter 1, Tannen's metamessage of speech was defined in terms of two interpersonal components, the speaker's *footing* and *affinity*, his wish for involvement or independence, closeness or distance. As a first point, with respect to *footing*, the above discussion revealed four possible ways of signaling unequal, i.e. inferior or superior, *footing* in English, German and possibly also Chinese: First, the overall direction of the nuclear pitch-patterns, which according to Tench (1995) can signal dominance (*falls*) and deference (*rises*); Second, the supralaryngeal *setting* of *palatalization*, which, based on evidence supplied by Laver (1980) and Ohala (1994), may contribute in the signaling of not only inferior physical size and strength but also an inferior, submissive, posture and attitude. As a third, feature, high overall pitch may also play a role in the signaling of inferior physical size and posture (Ohala 1994). Finally, the results of recent research in vocal emotion research, linking *breathiness* to the signaling of (positive) surprise and admiration, suggest that this phonatory *setting* may also have a part in the signaling of inferior speaker position. While the research findings relating to the overall direction of the English and German nuclear pitch-patterns, are not *a priori* applicable to the Chinese language, the results obtained for *palatalization*, *breathiness* and (high) overall may be tentatively applied to Chinese.

Coming now to equal *footing*, it is interesting to note that no suprasegmental form was found to be associable with this attitudinal message, a component of the interaction strategies *camaraderie* and *distance*. While this may strike one as surprising at first, the absence of a suprasegmental form signaling equal *footing* may account for two important phenomena, to which attention was called in Chapter 1: First, the strong ambiguity of *camaraderie* and *distance*, examined by Tannen (1984, 1986) and second, the phenomenon, also pointed out in Chapter 1, that listeners encountering these patterns of behavior often seem to be 'watching out', or 'listening for', the slightest possible sign of *superiority* in the speaker's voice. In fact, it is not surprising that such signs are often found, as there seem to be no suprasegmental phenomena serving to 'tip the balance', as it were, in the direct of equalness.

Turning now to the second component of Tannen's metamessage, *affinity*, the speaker's wish for closeness or distance, the results of the above discussion suggest that none of the suprasegmental forms is involved in the communication of this aspect of interpersonal information. It is thus concluded that a person's wish for closeness or distance may not be signaled by suprasegmental means.

In conclusion, a number of suprasegmental features may be associated with the signaling of inferior and superior *footing*, but neither equal *footing* nor the speaker's wish for nearness/closeness may be signaled by suprasegmental means. On the basis of these findings, two conclusions are possible: On the one hand, as none of the suprasegmental phenomena discussed in the literature are associable with the clear, unambiguous communication of both components of the metamessage of speech, it must be concluded that the communication of this aspect of interactional meaning is, as Tannen observes, a very different and uncertain affair. On the other hand, the obtained results may also be regarded as an indication that Tannen's association of the *metamessage* with the suprasegmental phenomena is erroneous, and that interaction strategies such as *camaraderie*, *distance*, *deference* and *support* are not signaled by suprasegmental means. However, the latter line of thought goes counter to the results of the many studies conducted in the ethnography of communication and interactional sociolinguistics and is therefore not supported here. Instead, we will continue to assume that *footing* and *affinity* are indeed signaled by suprasegmental means, but that further research may be necessary to determine the exact features responsible for these signaling processes. The experimental part of this work, to be discussed in Chapters 6 - 9, is devoted to finding out what these features are and how the metamessage, in particular the component of *affinity*, is communicated.

Part Three

Experimental work

The signaling of interpersonal attitude in German and Chinese

Introduction to Part Three

We now come to the experimental study conducted within the sociolinguistic and phonetic/phonological framework presented in Parts One and Two. The aim of the study was to determine in what manner suprasegmental phenomena are used in German and Chinese to communicate interpersonal attitudes of the kind discussed in Chapter 1, defined in terms of the interaction strategies *camaraderie*, *distance*, *deference* and *support*, and their antagonists *pushiness*, *cold aloofness*, *arrogance* and *false deference*¹. The experimental study focused on the interaction strategies *support* and *deference* and *arrogance* and *false deference*².

In preparation for the study, two minor alterations were undertaken with respect to the interactional components of *support* and *deference*, shown once more, with those of *camaraderie* and *distance*, in Fig. 44. First, given the fact that Tannen's 'wish for closeness and distance' is

<i>Camaraderie:</i>	<i>Wish for closeness/involvement and equal footing</i>
<i>Support:</i>	<i>Wish for closeness/involvement and unequal (superior) footing</i>
<i>Distance:</i>	<i>Wish for distance/independence and equal footing</i>
<i>Deference:</i>	<i>Wish for distance/independence and unequal (inferior) footing</i>

Fig. 44 The four interaction strategies with their interpersonal components (*affinity* and *footing*)

obviously not communicated by suprasegmental means³, the component of *affinity*, the degree to which a speaker associates, or identifies, himself with his interaction partner, was re-defined as follows: A person wishing to 'come close to' and 'associate himself with' another is likely to signal this with a warm, friendly undertone in his voice, while the voice of a person wishing to keep his distance may be expected to have a less warm, cool, possibly even cold undertone. Recent findings in vocal emotion research suggest a link between the undertone, or *valence* (described as 'positive' or 'negative') of a speaker's voice and phonatory voice quality (Scherer 1985, 1996). Tannen's epithets - the speaker's 'wish for nearness' and 'wish for distance' - were therefore replaced by the concepts 'warm' undertone or 'positive' *valence* and 'cool/cold' undertone or 'negative' *valence*, respectively.

As a second alteration, a change was made to the *affinity* component of *deference*. In Tannen's model *deference* involves the speaker's wish for independence/distance, corresponding to a cool/cold undertone and negative *valence* in the alternative model developed here. According to the observations of this writer, however, the communication of *deference* involves not a cool or cold, but rather a warm undertone. Therefore, the affinity component of *deference* was changed to a *warm* undertone/positive *valence*. The strategies *support* and *deference* and *arrogance* and *false deference* with their newly defined *affinity* and *footing* components, resulting from the above described alterations, are shown once more as follows:

<i>Support:</i>	<i>Positive valence and superior footing</i>
<i>Arrogance</i>	<i>Negative valence and superior footing</i>
<i>Deference:</i>	<i>Positive valence and inferior footing</i>
<i>False deference</i>	<i>Negative valence and inferior footing</i>

After this excursion we return to the discussion of the experimental study.

¹ It will be recalled from Part One that the antagonists of *camaraderie*, *support*, *distance* and *deference* are the interpersonal attitudes, or interaction strategies, in terms of which they are most commonly misunderstood (See summary of Part One). The numerous ways in which *deference* can be misunderstood are taken together here under the term *false deference*.

² This emphasis on *support* and *deference* was due to the importance of these two strategies for intercultural communication between China and the West. It was also motivated by the fact that misunderstandings of these two interaction strategies - in terms of *arrogance* and *falsehood/manipulativeness* - are perhaps most painful of all, as they involve serious distortions of speaker intent and misjudgements of speaker personality.

³ See conclusions of Part Two.

The experimental study was based on the following four questions:

- (1) Are the interaction strategies *camaraderie*, *support*, *distance* and *deference* - also found in German and Chinese conversational interactions?
- (2) Does the ambiguity of these interaction strategies, which was discussed above for the English language, also apply to German and Chinese.
- (3) In what way may the three groups of suprasegmental phenomena - intonation, the overall features of tempo, loudness, pitch and fluency, and voice quality - be found to contribute to the signaling of *support*, *deference*, *arrogance* and *false deference* in German and Chinese?
- (4) Are the suprasegmental features which contribute to the signaling of these interaction strategies responsible for their ambiguity and the misunderstandings of speaker attitude in conversational interaction between speakers of German and Chinese? What are the implications of the obtained results for intercultural communication?

To achieve the answer to these questions, dialogs with German and Chinese native speakers were recorded⁴ and the resulting speech material was subjected to a thorough speech signal analysis consisting of three stages. The answers to questions 1 and 2 were sought in the first two stages of the study, the sociolinguistic macroanalysis and the sociolinguistic judgement tests, presented in Chapters 6 and 7. Questions 3 and 4 were the objective of the phonetic microanalyses of the speech material, discussed in Chapter 8. The evaluation of the results, leading to the determination of the phonetic exponency of the interaction strategies, is the subject of Chapter 9. The final results of the experimental study and their implications for interpersonal and intercultural communication are presented in Chapter 10.

The methodology employed in this study was a combination of sociolinguistic and phonetic analytical procedures: In the first 2 stages the experimenter used the sociolinguistic methods of discourse analysis and judgement tests. At stage 3 of the study, phonetically oriented auditory and instrumental speech analyses were employed for the phonetic microanalysis. After the completion of the recordings, the recorded dialogs were first examined informally by the experimenter in the sociolinguistic macroanalysis. Using the sociolinguistic method of discourse analysis this analysis focused on the negotiation of the interactions by the German and Chinese interactants and the interaction strategies which they employed. Two kinds of passage were of interest: passages where the interaction strategies signaled by the speakers seemed clear and relatively easy to identify, and passages in which their speech behavior appeared ambiguous with respect to interaction strategy. Following this preliminary examination, judgement tests were organized with German and Chinese native speakers. Their task was to analyse utterances selected from the recordings in terms of the interaction strategy enacted in each case with the help of a questionnaire⁵. The tests focused on the interaction strategies of *support* and *deference* and their antagonists *arrogance* and *false deference*. Based on the results of the judgement tests, the analysed utterances were then classified according to type of interaction strategy. As a further category, all the utterances which had been judged to be ambiguous with respect to undertone, or *valence*, were grouped together under the heading 'ambiguous undertone/valence'⁶.

⁴ In keeping with the principles underlying Gumperz' methodology, no attempt was made to foresee the course of these interactions.

⁵ The organization and recording of the dialogs represent the 1st step of Gumperz' methodology, outlined in the Introduction to Part One. The sociolinguistic macroanalysis represent steps 2 and 3, and the judgement tests correspond to step 4 in Gumperz' methodology. Note that, as one difference to the procedure followed in Interactional Sociolinguistics, judgements of speaker attitude (interaction strategy) were based solely on the signaling value of the suprasegmental phenomena contained in the utterances. The original intent of speakers was not - indeed could not be - taken into account, in order to guarantee the objectivity of judgements.

⁶ The importance accorded to *valence* here was due to the fact, discussed above, that the interaction strategies *support* and *arrogance* and *genuine* and *false deference* differ first and foremost with respect to valence/undertone.

In stage three of this study, the classified German and Chinese utterances, or speech samples, each now representative of one or, in the case of ambiguous *valence*, two interaction strategies, were submitted to a phonetically oriented analysis of their suprasegmental features, using the phonetic methods of auditory tests and instrumental speech analysis. The results of these analyses were compiled in the form of a suprasegmental profile for each utterance, shown in Table 10 at the end of Chapter 8. The obtained suprasegmental profiles were then compared with the results of the sociolinguistic judgement tests in order to establish the phonetic exponency of the interaction strategies *support*, *deference*, *arrogance* and *false deference* in German and Chinese⁷. This is the subject of the discussions in Chapter 9, summarized once more in Chapter 10.

⁷ Admittedly, the treatment of speech behavior such as arrogance and false deference in terms of 'interaction strategies' may seem somewhat unusual, as it seems to imply that speakers actively adopt this kind of behavior, which is not normally the case. On the other hand, as was seen in Chapter 1, a person's originally intended interaction strategy is all too often (mis-) interpreted as representing this kind of behavior. It may therefore be concluded that at least in the minds of listeners, such 'strategies' certainly exist.

6. The interactional macroanalysis

The present chapter begins with a broad outline of the recorded dialogs, or *interactions*, giving a brief account of the recordings, the participants, their assignment, the negotiation of the interactions and the interactional events most frequently enacted by the participants. After this follows a comparative discussion of the four types of dialog, or *settings*⁸, which will focus on the characteristic manner in which these interactional events were negotiated by the German and Chinese participants, especially those referred to as the *speakers*. While the discussion will center on the differences and similarities between corresponding German and Chinese dialogs, it will also show that these dialogs represent genuine and naturalistic interpersonal interactions whose speech may serve as the basis for an analysis of the kind conducted here⁹. The occurrence of interruptions in the dialogs will also be examined briefly, as a matter of general interest.

6.1 The dialogs - a broad outline

6.1.1 The recordings

Two German and two Chinese women, hereafter referred to as the *speakers*, were asked to engage in conversational interactions with two men and two women each¹⁰. All the participants in the interactions were acquaintances of the experimenter who was thus familiar with their different personalities and conversational styles. The interactions were conducted in the form of role-plays, shopping-dialogs, using 'shopping lists' prepared by the experimenter. This familiar goal-oriented activity was chosen to serve as the outer frame of the dialogs in order to facilitate and channel the conversation between the interaction partners, as well as divert their attention from their speech, the object of the experiment. The recordings were made in the studio of the Phonetics Institute of the University of the Saarland, Saarbrücken, Germany, in July 1996, and took an average of ten minutes per dialog. Despite being sound-treated, the studio had a very pleasant and relaxed atmosphere, of great importance for the comfort and relaxedness of the interactants and the naturalness of the resulting dialogs. In order to ensure that the interactants would be quite at ease, they were left to themselves during the recordings, the experimenter leaving the room as soon as she had adjusted and switched on the recording instruments.

The speech signals and the laryngeal signals were recorded simultaneously with a Yamaha M8X multichannel tape recorder, the microphones being placed at a distance of approximately 30 cms. from the speakers' mouths. The laryngeal signals were recorded according to the procedure followed in electroglottography: Two electrodes, positioned on either side of the speaker's larynx, at the level of thyroid cartilage, and held in position by an adjustable neck-band, measure the electrical impedance between the vocal folds as a function of vocal fold contact¹¹. In the adjustment and fastening of the laryngographic neck-bands, great care was taken to ensure that the participants would not suffer any physical discomfort. Judging from the recordings, all the interactions proceeded in a smooth and relaxed manner. Only very occasionally did any of the speakers, mainly Clara and Sybille, show signs of nervousness or inhibition during the first one or two of their dialogs¹².

⁸ The ethnographic concept of the setting is extended here to include the type of interaction partner, i.e. familiar or unfamiliar male/female interaction partner, different in each of the four dialogs each speaker participated in.

⁹ As noted in the introduction to Part One, Gumperz does not object to the recording of constructed encounters as long as these are naturalistic speech situations in which the speech behavior of the interactants is not too strongly predetermined.

¹⁰ These four speakers were chosen, among other reasons, because they all had a large circle of friends who would be able to function as interaction partners and/or judges in the listening tests to follow.

¹¹ This is also referred to as laryngography or electroglottography. Discussions of the procedure and applications of electroglottography are found in Fourcin & Abberton 1971, 1977, Abberton & Fourcin 1972 and Fourcin 1981. See also Childers et al. 1984.

¹² Among the most frequent of such signs are speech disfluencies, i.e. pauses, hesitations, slips of the tongue and speech errors, mostly followed by pauses and restarts and unnatural sounding tones of voice. To name three examples, Sybille once says 'Rinderpack' instead of 'Rinderhack' (*minced beef*), Clara is once heard referring to 'the child' when she is actually speaking of her children, and the tone of

6.1.2 *The participants*

The recorded interactions involved two kinds of participant, the female speakers and their interaction partners. Since different degrees of familiarity between interaction partners give rise to differing interactional behavior and different types of interaction strategy, each speaker was familiar with one of her male and one of her female interaction partners, while the other two were strangers to her. The dialogs were all conducted in the native language of the speakers, i.e. Standard High German¹³ and Mandarin Chinese. As a rule, only native speakers of German or Chinese were chosen as interaction partners¹⁴.

The German speakers were Clara and Sybille, the Chinese speakers were Wu and You¹⁵. As the interactions evolved around the activity of shopping, they were chosen not only for their different personalities and conversational styles but also for their relative experience with the activities of shopping and cooking. Clara and You, in their early thirties, were lively, and full of humour, while Sybille and Wu, a little older, were more serious and composed. Of the two German women, both married with children, Sybille had more experience as a housewife than Clara, whose professional and political career limited the time she was able to spend at home with her family. Among the Chinese women Wu, married with one child, was the more experienced housewife. However, You, though a student and still single at the time, was an enthusiastic cook and an experienced and competent shopper, even in Germany.

The relative experience of the interaction partners in matters of shopping was also an important criterion in their selection, since this would be likely to influence the speech behavior of the speakers. The male and female interaction partners familiar with each of the four women were their husbands/boyfriends and woman friends of similar age and interests. While Clara's and You's partners were experienced shoppers, this was not the case with Sybille's and Wu's husbands. Clara's and Sybille's unknown female partner was a young student in her mid-twenties, just married and not yet a very experienced homemaker. The male interaction partners unknown to Clara and Sybille were two lecturers in their early 40s, very proficient in matters of shopping and cooking¹⁶. Since Sybille and Clara were good friends, they interacted with other as familiar female partners, and as You and Wu were from different parts of China and did not know each other they were asked to interact as unknown female partners with one another¹⁷. Their unknown male partners were a mainland Chinese and a Taiwanese student, respectively, both young men only moderately familiar with the activity of shopping¹⁸.

6.1.3 *The assignment*

Upon entering the studio the two interaction partners were asked to sit down at different tables opposite each other and the shopping/vending-lists were distributed. Unfamiliar interaction partners, who had waited for their turn in different rooms, were now introduced to each other¹⁹. In the interactions between unfamiliar interaction partners and familiar female partners, the participants were now asked to decide among themselves who would take on the

voice with which Clara thanks her husband for buying her flowers at one point is so unenthusiastic that it was interpreted as ironic in pilot listening tests. (This utterance was subsequently taken out of the speech corpus).

¹³ To be quite accurate, the two German speakers, who were from the Saarland, spoke the local variant of the High German Standard. Apart from minor intonational idiosyncrasies, discussed in Chapter 9, this does not differ substantially from other variants of High German.

¹⁴ As will be seen below, this was not the case in two dialogs. Though these interactions represent instances of intercultural communication, this will not receive special attention here.

¹⁵ The names of the speakers have been changed to conceal their identity.

¹⁶ These two lecturers, who were actually native speakers of Chinese and English, had been resident in Germany for over ten years and were quite fluent in German. They were chosen in their capacity as lecturers to ascertain that the atmosphere of the interaction remained formal and serious for as long as possible. For this reason also they were asked to attend the recordings in formal attire.

¹⁷ This combination of a Taiwanese with a Mainland Chinese speaker was meant to ensure that the interactants would be quite unfamiliar with each other and that the interactions would remain formal during all, or at least most, of their length.

¹⁸ A total of 14 participants were thus involved in the 14 recordings. As the speech of the 4 speakers was recorded and analysed separately, and two of the speakers served as interaction partners, this gave four different types of speech material for each speaker.

¹⁹ Different waiting-rooms were chosen for participants unknown to each other, so that they would not familiarize themselves with each other before the recordings. For the same reason, a screen was set up between unfamiliar interactants during the interactions.

role of the ‘vendor’ and who would be the ‘customer’. After this, they were asked to create a shopping dialog based on the items on the lists, it being stressed that they were to mention each item as often as possible²⁰. While they were free to improvise as much as they wished within the scope of this speech activity, it was however emphasized that they should remain natural, be themselves and not try to act a part²¹. In the interactions with the familiar men, the assignment was a little different: The participants were asked to imagine that the man had been out shopping with their girlfriend’s/wife’s shopping-list and was now about to ‘report back’ to her. The shopping/vending-lists were written in German and Chinese, respectively, and contained various kinds of fruit and vegetables, meat, sugar and other household staples, as well as miscellaneous items, such as music cassettes, ballpen cartridges and camera films. To make things a little easier for the participants, an approximate price-range was included for each article²².

6.1.4 *The negotiation of the interactions*

Though it may seem that the nature of the assignment in which the interactants were immediately involved - a role-play with a predetermined topic matter - should have limited their freedom of action and interaction and given rise to fourteen interactions whose course and content would be more or less identical, this was in fact not the case. On the contrary, due to the flexibility of the assignment with respect to the details of the speakers’ actions and speech behavior, each interaction took its own course and no two dialogs are alike. While the speakers’ roles alternated only between ‘vendor’, ‘customer’ and ‘wife’ and the general topic matter, relating to the items on the ‘shopping list’, remained largely constant, on the other hand, the details of the participants’ speech actions were not predetermined in any way at all: They were free to jointly negotiate each interaction according to their own ideas and preferences, this freedom involving not only the order in which the items were mentioned and the manner and speed with which they were discussed, but also extending to the speech actions, or *interactional events*, in which they engaged, such as greetings and farewells, enquiries as to the price, quality or weight of a given item and the pointing out of special offers. Most importantly, not only did the participants themselves determine what interactional events to enact, they were also free to choose the manner in which to negotiate them. As ‘customers’ they could request any kind of information or advice in the manner of their choice, praise or criticize the prices and decide whether they would purchase a given item, and as ‘vendors’ they could choose whether to recommend, praise or refuse the ‘customer’ certain articles, decide on what information or advice to give, and what ‘special offers’ they might have. Finally, as ‘wives’ they could also choose how to treat their men, in what manner to ask after the prices and whether to praise or scold them for their efforts. Hence, the interactants had total freedom with respect to the speech behavior they adopted in each setting. It was assumed that, rather than remaining constant in all the dialogs, the speech behavior of each speaker would vary both between and within individual settings, as they would be confronted with different interaction partners, whose personality and emotional disposition they would have to respond and adapt to²³.

²⁰ In this fashion it was ascertained that the items on the shopping/vending lists, which were identical in all cases, would recur a number of times during each interaction. Participants were allowed to introduce variants of articles, such as ‘spring onions’ or ‘preserving sugar’.

²¹ Pilot tests held prior to the recordings had shown that play-acting leads to largely identical and quite unnatural-sounding conversations. ‘Behaving naturally’ also meant that interactants familiar with each other used T-pronouns (in German with their appropriate verb forms), while participants unknown to each other were requested to use V-pronouns. The T-pronoun in Mandarin Chinese is *nǐ*³, the V-pronoun, which is only rarely used, is *nín*².

²² Similar to the procedure followed in the ‘map-tasks’ of experimental phonetics, where the maps of the two participants often contain different landmarks, the prices on the shopping/vending-lists were not always identical in order to create moments of confusion and potential conflict which had to be negotiated by the interactants. (Blakar 1975, Grice & Moosmüller 1995)

²³ What is referred to here as an interactional event is not to be confused with the ethnographic concept of the communicative event, briefly discussed in the introduction to Part One. The interactional event is used here to denote brief, more local instances of Gumperz’ speech activities.

Indeed, as the auditory analysis of the recordings showed, the manner in which the speakers negotiated each interaction varied strongly and in such a way as to suggest that both their emotional disposition and their attitude differed strongly both between and within individual encounters²⁴. The varying attitude of each speaker is clearly reflected in the different ways she negotiates the interactional events occurring in each dialog. To name one very conspicuous difference, while all conversations begin with a friendly greeting, the farewells at the end of the dialogs vary considerably in warmth, clearly reflecting the relationship between the interactants, which in half the dialogs had to be established in the course of the interaction. While in some cases the farewell is warm and cordial, the ‘vendor’ in one case even offering to help the ‘customer’ pack her shopping-bag, in others they seem cool, almost perfunctory, suggesting that the interactants met with problems and/or miscommunications in the course of their encounter²⁵. As another point contributing to the special character of each interaction, all the dialogs include a number of passages in which the interactants relate to different past, present or future real-life events, experiences, likes and/or dislikes, this probably being linked to the fact that they were asked to be themselves²⁶. A further result of this is that, although the dialogs were of course *staged*, in fact a very large portion of what the participants conversed about, whether this involved experiences, facts or feelings, was not invented at all but represented things as they really were in the lives and minds of the interactants. We now come to a discussion of the most frequently enacted interactional events and the manner in which they were negotiated.

6.1.5 *The interactional events*

The following ten kinds of interactional event are most frequently enacted in the interactions: On the part of the ‘customer’ and ‘wife’: (1) The requesting of information about the merchandise, (2) The voicing of special wishes for particular items, (3) Praise or criticism of the quality and/or prices of certain items; On the part of the ‘vendor’: (4) The enquiry as to what (further) items the ‘customer’ wishes to purchase, (5) The giving of information about the merchandise, (6) The agreement to/refusal of special wishes voiced by the ‘customer’, (7) Justifications given in response to criticism by the ‘customer’ or ‘wife’, (8) The pointing-out of special offers; (9) The enquiry as to what a given item is needed for, and (10) Discussions, mostly evolving around the use of certain items and relating to genuine real-life feelings and events. In the following, we will take a brief look at the five most important events.

6.1.5.1 *Asking for information*

As may be expected, the quality, quantity and the prices of the articles were of great interest to both the ‘customer’ and the ‘wife’. While enquiries as to the prices and the quantities are most frequent and at times quite detailed, as for instance when Sybille asks ‘vendor R’²⁷ to convert the price per kilogram of his kiwis to their price per piece, on the other hand, the issue of quality is also very popular and often leads on to longer discussions. The most frequently discussed topics are the freshness, ripeness and colour of fruit and vegetables, the country of origin of exotic produce, such as kiwis and oranges, and the provenance of the meat and sausages²⁸. As regards to spices, herbs, seeds and nuts, most questions relate to their consistency, i.e. whole, chopped, ground, cooked or raw etc., another very popular subject being

²⁴ While such changes in the attitude of the speakers could also have been brought about by other circumstances relating to the settings, precisely what produced these attitude variations was ultimately of no import for the experiment. The only thing that mattered was that different attitudes occurred and that they were identifiable in terms of interpersonal attitudes in the judgement tests.

²⁵ It should perhaps be pointed out that none of these interactional events were explicitly mentioned as part of the assignment.

²⁶ The genuineness of each such a passage was checked by the experimenter. The integration of the interactants’ genuine feelings and, especially, real-life mutual experiences is one of most important difference between these ‘naturalistic’ interactions and staged role-plays, where events are often invented and rarely shared by the interactants, giving rise to less genuine feelings and, it is to be expected, less natural speech behavior.

²⁷ The interaction partners are referred to with capital letters corresponding to the first, or first two, letters of their names.

²⁸ The provenance of the meat and sausages was important, as in the summer of 1996, the BSE crisis had just begun in Germany.

what they might be used for. Finally, enquiries about camera films, and ballpen cartridges focus on brand names and the number of exposures and cartridges per pack, while in the case of books and music cassettes, the most frequent questions relate to their titles and the names of the authors or artists.

6.1.5.2 *Special wishes and offers*

Special wishes and offers mostly relate to particular kinds of fruit, vegetables and meat, e.g. spring onions, soft kiwis, fresh, as opposed to frozen, duck and fresh *boudin* sausage²⁹, but also to certain baking ingredients, herbs and sweets, examples of this being Sybille's special offer of whole nutmeg, complete with a grater, both Clara's and Sybille's insistence on cooking chocolate as opposed to normal milk chocolate, and Clara's request for paper flowers and preserving sugar for making jam³⁰. In the Chinese interactions, special wishes and offers mostly pertain to particular ice cream or candy flavours, cigarette brands and types of dumping, i.e. with or without filling³¹.

6.1.5.3 *Refusals and offers of substitutes*

Refusals in which the 'vendor' states that a certain requested item is 'sold out' or 'out of season' are of particular interest since they go against the expressed wishes of the 'customer', are potentially face-damaging, and may thus be interpreted as indicating a potentially negative attitude on the part of the speaker. Some refusals are polite, even apologetic and followed by offers of other, similar items as substitutes, such as when Sybille, and later Clara also, explain that mangos are out of season and offer papayas and peaches in their place³². In a similarly friendly way, Wu apologizes to You that today she has no chicken eggs and no stewed chicken-wings, without however offering a substitute for the latter. In some cases, however, speaker attitude seems less friendly, as in the interaction between Sybille and her unfamiliar female interaction partner, where she offers no explanation and no substitutes for the absence of mangos, *boudin* sausage and flowers, which only a few minutes earlier she was quite happy to sell to her good friend Clara³³. Finally, in the rather special case of You, her refusals are delivered in a humorous, teasing and challenging manner, which are part of the teasing strategy of *tai-gang*³⁴ which she employs with her woman-friend as well as her unknown male interaction partner and which is meant to simultaneously frustrate and amuse.

6.1.5.4 *Discussions*

Discussions are found in every dialog, especially, but by no means only, in those between familiar interactants. In contrast to the discussions between husband and wife, which usually evolve around a conflict of some kind, those found in the conversations between woman-friends are friendly and harmonious and often accompanied by joking and shared laughter. These discussions typically begin with a request for assistance, information or advice from either one of the speakers, usually the 'customer', the answer to which then leads on to an often lengthy exchange of information and opinions. Typical examples are, for instance, Clara's and Sybille's discussion on how much preserving sugar to use for making what kind of jam, why Clara would want to decorate her house with paper flowers and which kind she should buy, and why the moonboots Sybille is selling have such a dark green colour. In the

²⁹ *Boudin* sausage is a special kind of sausage produced in France. Round and relatively thick, it is sold in slices.

³⁰ Neither spring onions, paper flowers, cooking chocolate, nor preserving sugar were included in the 'shopping-/vending-list'.

³¹ In Chinese this is *man-tou*. The choice between with *man-tou* with or without filling was also not mentioned in the 'shopping-list'.

³² The fact that mangos and *boudin* sausage were out of season in July was unknown to the experimenter and pointed out by Sybille. The peaches offered by Sybille as a substitute for the mangos also were not part of the 'shopping-/vending-list'.

³³ In a number of cases, the interpretations of the experimenter as to speaker attitude are tentative as they had to remain unconfirmed: The speech material was so sensitive and potentially face-damaging to the speaker that it was impossible to check back with them without seriously hurting their feelings. The accuracy of the answers received under these circumstances is also in doubt.

³⁴ *Tai-gang*, the challenging, teasing strategy typical of Chinese humour, introduced in Chapter 1, will be discussed in more detail below.

Chinese interactions, to which the above general points also apply, most discussions evolve around what a certain unknown food item is and how to prepare it, examples for such items being duck eggs, *spirit money*, avocados and *niu-nan* beef³⁵.

6.1.5.5 Criticism of the merchandise

Most instances of direct criticism, usually countered by justifications and/or excuses on the part of the interaction partner, are found in the conversations between the speakers and their husbands/boyfriends. Again, it is the quantity, the quality and the prices of articles which most frequently give rise to criticism, other reasons being their declared uselessness, unwholesomeness or the aversion of some family member to them, usually the children. In the German interactions a lot of the wife's critique is directed at the expensive kiwis, *boudin* sausage and ballpen cartridges, the unhealthy sweets and chocolate, and the potentially BSE - infected beef. In the Chinese interactions, it is the strongly smelling duck eggs, the unknown avocados, the *niu-nan* beef and the potentially unhygienic ready-made dumplings which meet with the most criticism. Wu's husband is also roundly scolded for having brought back *lichees* as well as *longan*, both fruit being of very similar consistency and taste - as well as very expensive³⁶. The excuses and justifications given by the husbands/boyfriends are diverse and often seem highly imaginative, although again, the large majority were not invented. The most frequent excuses are the cheap price, the excellent quality or some other favorable characteristic of the item in question. Thus, berated for buying a bag of exorbitantly expensive *peeled* potatoes, You's boyfriend explains that this will save her the trouble of peeling them. A little later, having been criticized for spending so much money on sweets, he retorts that this was necessary since You eats so many sweets herself that as a result they never have any left for visitors. As the price-ranges indicated on the husbands' and wives' shopping-lists often differed, there is also a large amount of confusion and controversy with respect to the prices of certain articles, which was resolved in different ways. Often, as in the interaction between Clara and her husband, the wife is simply told that she was mistaken, this kind of response however invariably leading to further discussion, since she has her own shopping-list, now a 'list of special offers', lying on the table before her³⁷. In some cases, the husband is criticized for having bought an item *not* featuring on the 'shopping-list' and as such judged to be unnecessary. Here, the husbands' most frequent reaction is to name some special or urgent need for the article in question, this kind of justification as a rule also referring to past or present real-life needs and events. Thus, criticized for having bought too large a pair of moonboots for his daughter, Clara's husband points out that by the time the winter comes, her feet will have grown to just that size. You's partner is also very good at this: Explaining why he bought a large, expensive, pack of ten blank cassettes, he reminds her of their noisy neighbor whom they have been planning to complain about for some time. Obviously, as he points out to her, they must first of all record the noise he makes as evidence before they can report him to the authorities³⁸. Finally, should the wife accuse the husband of having forgotten some item, his preferred way out of the crisis is to declare that it is not needed or indeed disliked, either by himself or another family member, especially the children.

The aim of the foregoing discussion was to show that, despite the fixed outer frame and the predetermined topic matter of the dialogs, the participants did not simply act out their roles as 'vendor', 'customer' and 'wife', inventing everything as they went along, but engaged in what

³⁵ While duck eggs and *lu-rou* (a kind of stewed meat) are well-known in Mainland China, *spirit money* (*ming-zhi*) and *niu-nan* caused a certain amount of confusion among the participants from China. While *spirit money*, used at religious ceremonies, is slowly making a comeback, the term *niu-nan*, denoting a certain cut of beef, may be limited to Taiwan. It appears that the corresponding term in China is *niu-du*. As avocados (*luo-li*) have only recently been introduced to China and Taiwan, they also incurred a lot of confusion with both the Chinese and the Taiwanese participants. Mostly, they were thought to represent a kind of pear (*li*).

³⁶ The *longan*, literally dragon's eye, is a small round fruit whose appearance and taste strongly resembles that of *lichees*.

³⁷ This was checked with the participants after the recordings.

³⁸ The problem with the noisy neighbor was real, H's idea with the cassettes was spontaneous.

may be termed quasi-natural, or naturalistic, conversational interaction, in the sense that the manner in which they spoke and interacted and the feelings they displayed were genuine and corresponded to the way they spoke, thought and felt in real life. True to the assignment, the interactants 'mapped' themselves - their personality, their way of thinking and their conversational style - 'onto' the 'vendors', 'customers' and 'wives' whom they were portraying, and their varying feelings and attitudes gave rise to different ways of negotiating the interactional events of each encounter, many of which occurred in each of the four settings.

The different ways in which the speakers negotiated the interactional events, reflecting the different feelings and attitudes which they entertained in each case, are interpretable in terms of different interaction strategies whose metamessages communicate these attitudes³⁹. Thus, for example, consenting to provide information or granting the 'customer' a special bargain in a warm and friendly manner is representative of warm, kind and supportive behavior, while the refusal of a requested item or strong, unfriendly criticism of a purchased item is strongly indicative of a superior and negative, i.e. arrogant, attitude. On the other hand, soliciting the interaction partner's assistance and advice in a humble and inferior manner may be interpreted in terms of deferential behavior, either genuine or false. As the next point we will now take a look at the ways in which the interactional events were negotiated by the speakers.

6.2 *The negotiation of the interactional events*

The comparative analysis of the German and Chinese dialogs revealed two interesting points: First, apart from small differences possibly reflecting slight differences in speaker attitude, the manner in which the four speakers enact and negotiate the interactional events of one type of setting⁴⁰ is very similar, so that each of the four settings is describable in terms of a relatively small set of interaction strategies. Second, apart from very few minor differences which are mostly differences in conversational style⁴¹, the interaction strategies employed by the four German and Chinese speakers in one type of setting are also almost identical and communicate much the same underlying feeling(s) and speaker attitude(s). In the following discussion of the four types of interaction we will discuss the interaction strategies which appear to be typical for each setting and the attitude(s) of the speakers signaled by their use.

6.2.1 *The interactions with familiar women*

The interactions between the speakers and their woman-friends are marked by an atmosphere of warm friendship, merriness and a high degree of harmony, this no doubt due to the fact that the interactants knew each other well and were familiar with each other's lives and habits, preferences and dislikes and their conversational styles. Due to their familiarity with one another, it may also be assumed that the power relationships between them were firmly established so that the power battles sometimes found in other *settings* are conspicuously absent here. As both women were able to draw upon the common background knowledge they shared, the course of their conversations is also very smooth. Throughout the dialogs their speech behavior mirrors the friendship that binds them in real life, often reflecting subtle traits of their relationship. Interestingly, simultaneous speech is very frequent and there are also a large number of interruptions in all three dialogs⁴². None of these, however, appear to have

³⁹ The four types of interaction, or setting, are defined according to the identity of the interaction partner: (1) unknown man, (2) familiar man (3) unknown woman (4) woman friend.

⁴⁰ As observed above, the varying attitude of a speaker was not necessarily always of a truly interpersonal nature but may sometimes have reflected her attitude with respect to some other aspect of the interaction. However, since in all cases speaker attitude was indeed interpretable in terms of a certain interpersonal attitude, in the following 'attitude' will be taken to refer to 'interpersonal attitude'.

⁴¹ In some cases these also reflect differences in culturally-patterned communicative conventions, relating to the norms, or systems, of politeness discussed in Chapter 1.

⁴² Defining what exactly is to be regarded as an interruption is a matter of some dispute. While some researchers, following Schegloff et al. (1973), regard all instances of simultaneous speech as interruptions, others, such as Tannen (1984, 1986 and 1993b) call for a distinction between truly disruptive simultaneous speech and supportive simultaneous speech which aims to signal support, interest and sympathy to

been of a disruptive nature. On the contrary, in the vast majority of cases, the speech of the second speaker communicates camaraderie and support by expressing her sympathy and understanding for the first speaker's wishes and feelings, contributing to the atmosphere of harmony, warmth and friendship⁴³. Characteristic of these interactions is thus speech behavior serving to signal friendship and harmony, with details depending on the character and the *footing* of the interactants. Depending on whether the speaker occupies a superior or inferior *footing*⁴⁴, the interaction strategies most commonly employed here are *camaraderie*, *support* and *deference*. This is illustrated in the following excerpts⁴⁵. In 'Moonboots', Clara's complaint as to the colour of the moonboots, in lines 5 and 7, is dealt with by Sybille (in line 6) in a friendly, perhaps slightly superior but warm manner (*support*), while in 'No Mangos' Clara's disappointment at the unavailability of mangos in line 6 (as example of *deference*) is met with Sybille's suggestion of a substitute, delivered in a very warm, almost motherly manner (signaling *support*). In the next excerpt, 'Inviting guests', two passages are of interest: In line 2 partner's S's allusion to Wu's predilection for chillis represents an example of the way in which common background knowledge is built into the conversation (*camaraderie*), and in lines 1, 5 and 11, Wu's gently prompting questions show her reaction to the initial ner-

'Moonboots'

- C: (1) I noticed + you've got *moonboots* on special offer. (mid-slow, sweet, childlike)
 S: (2) Yes, *autumn's* on its way, you *know*. (mid-fast, warm)
 But these are only moonboots for *children*. [Up to size thirty-seven, that's *all*.
 C: (3) [Only moonboots for *children*.. (slow, pensive)
 Okay, I guess that'll be all *right*. What *colours* are they then? (mid-fast,
 S: (4) They're usually *black*, with, sort of, colourful *patterns* on them. warm, but matter-of-fact)
 Like, here we have a *black* pair with + with blue *zigzag* patterns.
 What *size* do you need then?
 C: (5) No, Sybille + I'm afraid I don't think I like *that* kind of colour (slow, a little softer,
 on *moonboots*.... disappointed)
 S: (6) Well, they're usually not very *light* because otherwise they look really *dirty*, (brisk,
 you know, in *winter*. matter-of-fact)
 C: (7) Yes, but all the *same* ++ No, I *don't* think... [No. (as before)
 S: (8) [Mm [So, none of these moonboots is for *you*, Mmm (soft, sympathetic, motherly)
 C: (9) Moonboots ++ I'm afraid *not*. They looked + they looked prettier in
 the *leaflet* (sweet, childlike)
 S:(10) Okay (warm, accepting)

the first speaker. A special case of the second type is the back-channeling signal, discussed in passing above. Three criteria were employed here for a broad distinction between disruptive and non-disruptive simultaneous speech: (1) timing, (2) the effect on the first speaker, and (3) the semantic relation between the speech of the first and speaker. With respect to timing, an interruption was defined as occurring well before what is referred to as a TRC in the literature - a point at which a change of speaker is likely to occur - and before the first speaker has given any indication that he is about to complete his turn. With respect to the second criterion, an interruption was defined as incurring a disruption in the first speaker's flow of speech. And, finally, the semantic content of an interruption is generally *atypical*, i.e. outside the immediate topic-matter treated in the original speaker's contribution. While a further possibility would have been to examine whether original speakers showed any signs of confusion or annoyance caused by a potential interruption, this possibility was limited in these dialogs, as they were not video-recorded. No *verbal* expressions of confusion or annoyance occurred, however, in any of the interactions. In the following transcriptions, hyphens mark the points at which interruptions, defined in this manner, occur.

⁴³ According to some researchers working in Language & Gender, non-interruptive, supportive, simultaneous speech is more typical of female speech. In contrast to this, male speech seems less supportive, more strongly competitive, and simultaneous speech is of a more aggressive nature, aiming to challenge or 'attack' the current speaker. See Leet-Pellegrini 1980, Maltz & Borker 1982, Tannen 1990, the papers collected in Coates & Cameron 1989 and in Coates 1998.

⁴⁴ Presumably due to the mapping of the speakers' personalities onto their roles, the *footing* of a speaker in a given setting was found to be quite independent of the role she was enacting. Instead, the *footing* of a speaker depended on two factors: First, on her proficiency in matters of shopping and cooking, as compared to her interaction partner, and secondly, on the communicative function of her utterance in each case. When seeking help or asking for advice, for example, a speaker normally takes on an inferior *footing* with respect to the listener, whose goodwill/knowledge she is appealing/deferring to.

⁴⁵ In the translation of these passages, whose original versions are to be found in Appendix I, Section 3, free translations were chosen in order to faithfully reproduce their overall feeling and the interpersonal attitude signaled in each case. The symbols used in the transcription of these passages are presented above, at the end of the General Introduction. As these passages contain translated speech material, there is no notation of intonation. However, overall pitch, loudness and tempo as well as voice quality are indicated, the latter using impressionistic terms defined below. Sentences ending with small dots have very pronounced final slow-down and sound very unsure.

vousness of her friend, guiding her through the interaction in a very gentle, considerate way (*support*).

'No Mangos'

- C: (1) Now, I need a few things from the + er + *fruit* corner + *Mangos, Bananas* +++ and *Kiwis*. (sweet, childlike)
 S: (2) [Mm
 C: (3) How much are the *mangos*? (mid-fast, lively)
 S: (4) The *mangos*, they + I'm afraid we don't *have* any at the *moment*. (slow, soft, gentle)
 Do you need them *urgently*?
 C: (5) [Ohh! (very disappointed)
 C: (6) Oh + in this *recipe*, it said I needed one mang-
 one mango + one to two *mangos*... (slow, disjointed, distressed, almost tearful)
 S: (7) Hm, mango is a *problem*. But I *do* have *papayas*, if you *like*. (slow, warm, sympathetic, motherly)
 C: (8) Yes, *okay*.. (a little happier)
 S: (9) They cost two marks ninety *each* and really taste quite *similar*. (as before)
 C: (10) Yes, *okay*...
 S: (11) I'm sure you can use them *instead*. (as before)
 C: (12) Mm.Mm. (relieved, happy)

'Inviting guests'

- W: (1) I have *guests* coming *tomorrow*, so I want to buy a little more than *usual* today. (mid-fast, warm, friendly)
 Let me + Let me first take a look at the *vegetables*.
 What *vegetables* do you *have* here? (slow, soft, warm, gently coaxing)
 [0.8]
 I see you have + *chilli* peppers, *right*? (warm, gently coaxing)
 S: (2) [We have [chilli peppers
 You *like* chilli peppers, *don't* you ? (slow, soft, hesitating, very unsure)
 W: (3) Yes, I *like* chilli peppers, yes yes (as before)
 S: (4) [Mm [Mm
 W: (5) Then, you also + have *onions*, *right*? (again gently helping S on)
 S: (6) [Mm, Mm..
 We have + *onions* + *pumpkins* for. making *soup*. (as before, still rather unsure)
 W: (7) Ah yes, *pumpkin* is *great*! (Enthusiastic) Do you have any *other* *vegetables*? +++ Er (fast, enthusiastic, louder)
 S: (8) Er, yes, we also have + *potatoes*. (louder, faster, becoming a little more confident)
 W: (9) Ah yes, *potatoes* right. That's a *good* idea, a *potato* soup. (as before)
 S: (10) [Mm, Mm
 W: (11) Mmm. Now, on the other hand + *fruit* + I suppose I should get some *fruit*,
 too, *shouldn't* I? (slow, soft, again prompting S)
 S: (12) Mm, yes, of *course*. (as before)
 W: (13) The fruit do you *have* here...
 S: (14) For example, right now, we have fresh *lichees*, [and *longan* (mid-loud, mid-fast, confident)
 W: (15) [Ahh, *lichees*, yes
Lichees, longan, yes...
 S: (16) [Mm

While joking and shared laughter is found in all these dialogs, the mild, gentle kind of joking typical of the interactions between Clara and Sybille (German) and Wu and S (China) contrasts strongly with the teasing and challenging humour, *tai-gang*, found in the conversation between You and her woman friend F. *Tai-gang* is a feature of You's conversational style and her use of it is shown in the following two excerpts, '*Orchids*' and '*Only eight Marks*'. In '*Orchids*', after having first enthusiastically offered F three different kinds of orchid to select from (line 2), You suddenly changes course in the same breath and points out that she actually only has one kind, so that in effect F has no choice at all. Similarly, in '*Only Eight Marks*', after having stated in line (2) that lichees are in season now and therefore cheap, You goes on to name the exorbitantly expensive price of eight Marks per kilogram, which is so unexpected that F is first rendered speechless, then bursts out laughing. When F goes on to protest that this in fact very expensive, You rebuts this, pointing out that here in Germany, such a price for fruit from China cannot possibly be considered expensive. F's

following decision to forget about the lichees and buy plums instead (line 5) does not help her much, as these turn out to be even more expensive. This is doubly unexpected as the plums are ‘made in Germany’ and You has pointed out how cheap they are right now (line 6). In both cases, You’s use of *tai-gang* is meant to simultaneously frustrate and amuse F, as well as challenge her to ‘join in the game’ by countering with equally funny, but at the same time frustrating and potentially face-damaging remarks. Interestingly, in this interaction F, who is good at *tai-gang*, did not join in but always gave in in a gentle and deferent manner.

‘Orchids’

- F: (1) Now..... I also need two pots of orchids. (mid-slow, warm, friendly)
 Y: (2) Orchids, okay, two pots of orchids. Would you like butterfly orchids⁴⁶?
 Or slipper orchids? Or rather New Year’s orchids? Or
 Do you have any- any special wishes concerning orchids? (mid-loud, fast, enthusiastic,
 However, I’m so sorry, we only have butterfly orchids! challenging, mock serious)
 F: (3) (Laughs) Oh, okay. Well, I guess I’ll take butterfly orchids then!

‘Only eight Marks’

- F: (1) Well, what about lichees then? (mid-soft, a little wary)
 Y: (2) Lichees, ah yes. Er + lichees have just +++ come into season. So they’re, (mid-loud, fast,
 er, really cheap right now. One kilogram only costs eight marks! (excited, enthusiastic)
 [0.8]
 F: (3) Only - eight - marks?! (laughs) What do you mean?! That’s really expensive! (mid-loud, slow, incredulous)
 Y: (4) Right! (as before)
 Y: (4) Absolutely not. At this time of year, and in Germany, too. Eight marks per kilogram
 is very cheap! (mid-loud, fast, challenging, mock serious)
 F: (5) Oh, well in that case, I think I’ll take plums instead. (smiles)⁴⁷ (soft, mid-slow)
 Y: (6) Good, plums. Ahhh, plums are very cheap right now! One kilo
 only costs nine marks! (as before)
 F: (7) Only - nine - marks?! (mid-loud, slow, laughs)
 Y: (8) Right! (mid-loud, challenging, as above)

To summarize, the interactions between woman friends are characterized by a cheerful and warm atmosphere and speech behavior communicating goodwill, friendship and harmony. The interaction strategies typical for these interactions are *camaraderie*, *support* and (genuine) *deference* which in the Taiwanese dialogs alternate with *tai-gang* humour. While the true metamessage of *tai-gang* is one of closeness and mutual friendship, on the surface it involves the display of (playfully meant) *arrogance* which momentarily places the hearer in an inferior position. The hearer then counters with the same behavior, latching onto⁴⁸ the last remark of the first speaker, who counters once more. This ‘teasing-match’ continues in the same manner until one of the interactants runs out of ammunition, so to say, and ‘surrenders’. We will return to *tai-gang* below, in the interaction between You and her male interaction partners.

6.2.2 The interactions with unknown women

Compared to the interactions between woman friends, the dialogs between women meeting for the first time are less warm. Instead, the atmosphere is one of polite distance, with warm and cool, at times cold, passages alternating with each other. As these women know nothing about one another, a great deal of their conversational work aims at feeling their way towards each other. Since at the same time they must also complete their task, the shopping-dialog, it is not surprising that small errors and mishaps occur quite often in these interactions. It is also

⁴⁶ Note that in this interaction You addresses her friend with the polite form of address reserved (*nin*) for situations of inequality/power, which is really against the rules. However, both young women are still enacting their true personalities.

⁴⁷ Due to her familiarity with the speakers’ voices, the experimenter was generally able to discern with some accuracy the presence of *smiling*, reflected in the auditory features of *labiodentalized voice*, in the speech of the interactants.

⁴⁸ The term ‘latching-on’ refers to utterances of different speakers which follow after one another without any pause (Tannen 1984).

often noticeable from the subtle details of the interactants' behavior that many of these errors and mishaps, which usually involve no more than the wrong choice of words or a slightly too confident - or too submissive - sounding tone of voice, are interpreted in terms of an unwelcome attitude by the other speaker. Though no open conflicts occur, the two women often engage in what may be termed as 'power battles', each trying to impress and outdo the other with her knowledge of food, prices and cooking. This is particularly apparent in the interactions between S and Wi, two German women with entirely different lives and personalities. While the initial atmosphere in their dialog may be referred to as one of friendly distance, the general feeling slowly deteriorates and finishes with a cool, perfunctory farewell. In the interaction between You (Taiwan) and Wu (China), things are only marginally better. Both women have a friendly, flexible and highly tolerant nature and they seem to have brought a large amount of goodwill into the interaction. However, their tolerance and goodwill is repeatedly put to the test by the large number of mishaps which occur, relating both to the language and the different cultural backgrounds of the two women, who also have somewhat different conversational styles⁴⁹. Despite recurring efforts to repair the damage, this interaction has a certain number of distinctly cooler passages and ends without a farewell. Interestingly, in stark contrast to the other two, the interaction between C and Wi is characterized by an atmosphere of friendly distance, alternating with warmth. Although errors also occur here, in all cases repair is swift and clearly more effective than in the other dialogs. There are no critical moments and the dialog ends with a cordial farewell, C offering Wi an additional carrier-bag for her purchases. With respect to simultaneous speech, it is interesting to note that while the cool interactions between Sybille and Wi, and Wu and You, contain very few instances of overlap, the warmer dialog between Clara and Wi has a very large number of simultaneous turns and interruptions, which are, however, clearly of the same harmonious nature as those encountered in the interactions between familiar women. It seems that the scarcity of simultaneous speech in the other two dialogs reflects the very careful speech behavior of the women, who were trying to keep a certain distance from each other. While this may be interpreted as reflecting a cool and distant interpersonal attitude, their reticence may also have been due to their unfamiliarity with each other's conversational styles, including matters of tempo and timing. In this case, the scarcity of simultaneous speech reflects the women's efforts not to verbally 'step on each other's toes', i.e. interrupt each other.

Thus, generally speaking, these interactions contain (at times slightly false-sounding) *deference* and (possibly slightly cool) *support* and there is a very sporadic incidence of *camaraderie*. A rather large number of utterances sound suspiciously superior and cool, i.e. *arrogant*, and also seem to have been interpreted as such by the interaction partners. The friendly atmosphere between C and Wi is illustrated in the excerpts 'Summer onions' and 'Mangos'. In 'Summer onions' Clara's mistaken reference to spring onions as 'summer onions' (in line 2) is corrected by Wi in a friendly manner in the following line. Clara's reaction to this correction, potentially face-damaging as it calls attention to her non-proficiency in matters of shopping, is equally friendly and almost apologetic. In 'Mangos' Clara admits very apologetically in line 5 that she can't remember what Wi wanted to buy next⁵⁰. Her suggestion that it was mangos turns out to be correct. In lines 7 - 9 it then emerges that these are very expensive, almost too expensive for Wi to afford. While Clara sticks to the price (in line 9) and does not seem prepared to bargain, she does signal her sympathy by sighing apologetically and explaining

⁴⁹ In China and Taiwan slightly different forms of the Mandarin Chinese standard are spoken, a fact which is reflected in the terms used to denote the official standard language on both sides of the Taiwan Strait - *guo²-yii³* in Taiwan and *pu³-tong¹-hua²* on the Chinese mainland (Norman 1988: 135 f.). One example of the differences between the two standards - the Taiwanese word *niu-nan* - as compared to the Mainland Chinese *niu-du* - was noted in passing above. To some degree, therefore, interaction between native speakers of *guo²-yii³* (Taiwan) and *pu³-tong¹-hua²* (Chinese Mainland) may be referred to as intercultural communication.

⁵⁰ A few minutes prior to this, Wi had mentioned a few things she intended to buy, including the mangos referred to here.

that mangos are out of season. In all of Clara's utterances here, the feeling which emanates from her gentle, breathy and childlike tone of voice and hesitating, disjointed manner of speaking is one of inferiority and *deference*, which, given the fact that she is in the role of the 'vendor' is rather striking⁵¹.

'Summer Onions'

- Wi: (1) Now then, I *also* need +++ three pounds of onions. (mid-fast, friendly)
 C: (2) Three pounds of *onions*. Would you + May I recommend you these *summer* (sweet, childlike, disjointed, seemingly unsure)
 onions with- er, on-, with the *stems*? (matter-of-fact, but friendly)
 Wi: (3) You + you mean the *spring* onions? (sweet, friendly, almost apologetic)
 C: (4) Yes, these *spring* onions.

'Mangos'

- C: (1) Then you said you wanted *oranges*... (mid-fast, sweet, friendly)
 Wi: (2) [Yes.
 C: (3) Here, ++ er + how much oranges would you *like*? (as before)
 Wi: (4) Two *pounds*. (as before)
 C: (5) Two pounds of *oranges*, *here* you are! And, er ++
 I'm awfully *sorry*, but I've completely *forgotten*. I think (soft, slow, unsure, apologetic)
 mangos was what you wanted *next*, wasn't it?
 Wi: (6) [Yes. [How much are the *mangos*?
 C: (7) I'm afraid they're a +++ bit *expensive* right now. (slow, unsure, childlike, apologetic)
 Two mangos cost eleven marks *twenty*. (a little cooler, a little displeased)
 Wi: (8) Two *mangos*. And what if I only buy *one* mango? (soft, warm)
 C: (9) [2.5] Hm, (sighs sympathetically) then that'll cost you *half* as much. (cool, disappointed, but not unfriendly)
 Wi: (10) Well, I *do* need them, I mean, + for *cooking*
 C: (11) [Okay. [Yes
 Yes, of *course*, I mean, if the recipe *says* so... I really am very *sorry*, (slow, sympathetic, disjointed, seemingly helpless)
 but it's just not the right *sea*- the right time of the *year* for mangos... (a little warmer)
 Wi: (12) Yes, well, I suppose I'll have to *accept* that, I definitely need a *mango*.
 C: (13) [Mm, Mm

In the next excerpt, 'Whole goose', taken from the interaction between Sybille and her unknown female interaction partner, also Wi, the atmosphere, which begins quite friendly, slowly deteriorates, as neither of the women is quite sure how to interpret certain utterances of her interaction partner. The process starts with Sybille's very emphatic question in line 3, which may have sounded a little too surprised and exaggerated, and therefore possibly a little accusatory. Wi manages to go along with this, but in line 6, when consenting to wait for her goose until the next morning, she points out quite emphatically that 'it is rather tight'. Her strong, insistent and thus slightly accusatory undertone is underlined further by Wi's following remark (in line 12), where she hints that she may decide to buy the goose somewhere else. Finally, in line 14 where she asks Sybille to order the goose, her request has a somewhat 'gracious' and superior undertone, due to the matter-of-fact tone of voice and her choice of vocabulary. The possibility that it was interpreted as such by Sybille is shown in her next turn, where she seems to be pointing out in rather cool manner that Wi should perhaps be thinking of baking something for her party⁵².

'Whole Goose'

- Wi: (1) Then I won't buy any *boudin* sausage. But I do need a *goose*. (mid-fast, matter-of-fact)
 S: (2) [Mm
 S: (3) A *goose*? [You mean] a *whole* *goose*? (surprised, emphatic)
 Wi: (4) [Mm [Yes, a *whole* goose, yes. (as before)
 S: (5) In *that* case I'll have to *enquire* first. One moment, *please*. (still quite friendly)
 [1.5]⁵³

⁵¹ This is therefore a good example of the fact, mentioned above, that the interactants truly mapped their own personalities onto the roles they were asked to play.

⁵² Note that up to this point Wi had not mentioned baking before, so that Sybille is not merely reminding her of it, but pointing it out.

⁵³ During this period of silence, Sybille pretends to leave the counter to enquire about the availability of fresh geese.

- Okay, right now we don't *have* any. But, if you asked *again*, say, (friendly, but matter of fact)
tomorrow *morning*? Would that be *okay* for you?
- Wi: (6) Tomorrow *morning*? That's a *bit tight*, but... (mid-loud, very emphatic)
S: (7) Well, I can *try*, but at the moment we don't *have* any. Perhaps (a little cooler)
you could give-
- Wi: (8) [Will it be *fresh* tomorrow *morning* then? (matter of fact)
S: (9) [Yes, It'll definitely be *fresh* then. (slow, pensive)
Wi: (10) Hm, it'll be *fresh* then. (friendlier)
S: (11) Perhaps you could give me a *call* this *afternoon*, I'm sure I'll know (slow, pensive, slightly cool)
for sure *then*. (cooler)
Wi: (12) Okay, yes. Otherwise, I *could* of course try somewhere *else*. (mid-fast, matter-of-fact)
S: (13) [Mm, yes. (cool)
Wi: (14) Yes. Okay, you can go *ahead* and order the *goose* for me.
S: (15) Mm. +++ Aren't you planning to *bake* anything for your *party*?

The distant, at times decidedly cool atmosphere in the interaction between You and Wu is shown in the following excerpt 'Avocado'. As noted above, the avocados (*luo-li*) caused a high degree of confusion in a number of the Chinese interactions because only very few of the participants knew what they were. In this excerpt Wu, who is convinced that *luo-li* are a kind of pear (*li*) wants to know more about their taste and consistency, so that she can decide whether to buy *luo-li* in addition to the (Chinese) pears she has already purchased. Her first question (line 1) is however misunderstood by You, who believes she is asking about the price. After Wu realizes her mistake, she tries again, in line 5, where she also adds her reasons for wanting to know. Interestingly, You, a very warm and lively young woman, seems a little unwilling to provide Wu with the information she needs. Her answers in line 6 and 7 are reticent and not truly informative, and seem rather cool, especially as she does not react to Wu's apologetic laughter (in line 7) by laughing or even smiling (in line 8). The contrast between the speech behavior of the two women - sweet, respectful and friendly on the one hand and rather cool on the other - is underscored by Wu's hesitating and disjointed way of speaking and her humble, inferior manner which is clearly meant to signal *deference* - the *deference* of one not knowing and seeking information from one who is assumed to know. On the other hand, it could be that it was this humble, inferior manner which 'turned off' You, who may have interpreted it as *false deference*⁵⁴.

'Avocado'

- W: (1) Now then, this *avocado* what kind of *fruit* is it? + Is it (friendly, slow, soft,
I mean, is it similar to *pears*? hesitating, disjointed)
Y: (2) Oh, *no*, not at *all*. (mid-loud, matter-of-fact, a little cool)
W: (3) Not at *all*, I *see*. [Avocado...
Y: (4) Mm. [Avocados are + twelve marks per *kilogram*.
W: (5) Hm, +. *twelve* Mar- So, compared to *pears* Which- which of them tastes *better*? (fast, soft, disjointed, unsure)
You *see*, I want to *compare* them a little, I only want to buy *one* kind.
Y: (6) That + depends on one's *taste*. (mid-loud, matter-of-fact, a little cool)
W: (7) That depends on one's *taste*. Well, which- which kind is *sweeter*? (as before)
I like *sweet* fruit, you *see*. (laughs apologetically)
Y: (8) Oh, avocado is *not* very sweet (doesn't laugh or smile) (as before)
W: (9) Oh, I *see*. Avocado is *not* very sweet I *see*...

To conclude, the underlying feeling of these interactions is one of polite distance, cool undertones alternating here and there with more friendly passages. Although the interactants also employ the strategies of *deference* and *support*, especially in the interactions between Sybille and Wi and Yang and Wu these often sound less warm and therefore less sincere than those employed in the dialogs between women-friends. With the possible exception of the dialog

⁵⁴ It is possible that You was unable to decide whether Wu was simulating this ignorance or whether she really did not know, the latter of which, given Wu's obvious seniority, may have seemed rather unlikely to her. This passage was one of those considered too sensitive to ask the participants' opinion about.

between Clara and Wi, there is no allusion to mutual wishes, likes and dislikes (*camaraderie*) and a high incidence of utterances with a distinctly cool and superior undertone (*arrogance*).

6.2.3 The interactions with familiar men

In these conversations, the women's attitude may be broadly described as suspicious and/or negative. In all but one of these interactions, the men are treated in a very superior, often extremely unfriendly manner, the speakers' superior *footing* due to the fact that as women they know more about food and prices, no matter how often the husband has replaced them doing the weekly shopping. From the beginning, all four 'wives' display a (very) suspicious, often negative attitude regarding the result of the 'husband's' shopping expedition⁵⁵. Beginning with a very short, perfunctory, sometimes quite unfriendly greeting on the part of the woman, all four dialogs develop almost immediately into a kind of *crossexamination*, with the husband being told to report on what he has bought and how much every item cost. While there are passages in which both partners discuss certain topics as equals, during much of the time the husband is subjected to very face-damaging treatment, almost everything he says being criticized. Interestingly, all four men go along with this treatment, a fact which is particularly surprising in the Chinese dialogs, giving the traditionally inferior position of women in Chinese society⁵⁶. While Wu's husband Ch. does protest here and there, it is a mild, coaxing and at times amused protest, and You's boyfriend H. remains quite calm despite the strong and ironic, at times almost sarcastic, criticism which You hurls at him. The fact that many if not all of You's attacks may have been meant as *tai-gang* is hinted at by the fact that H. 'shoots back' very well in the second half of the interaction. On the other hand, there is no indication at all that Wu's scoldings are part of this joking strategy, especially as *tai-gang* is not part of her conversational style. Not surprisingly, praise is very rare in these dialogs: Sybille's husband is praised once, Clara's twice, Wu's husband three times, You's boyfriend not at all. Coming now to simultaneous speech, this is relatively frequent in these interactions, especially in the Chinese dialogs, where the large number of interruptions is particularly conspicuous. Even more noteworthy than their number, however, is the character of the overlaps found in these settings: They are much more negative and disruptive than those encountered in the dialogs between women. While in the German interactions supportive and disruptive overlaps alternate, the dialogs between Wu and Ch. and Yang and H. are marked by the almost exclusive presence of simultaneous speech in which the speaker contradicts, calls into question or makes fun of her partner's preceding or ongoing contribution or expresses her impatience or displeasure with him. The most frequently encountered inter-action strategy in these interactions is cool superiority (*arrogance*), sometimes possibly part of the *tai-gang* strategy. Here and there, especially in the interaction between Clara and her husband H, there are also a number of utterances which seem to represent the strategy of *deference*, and there are a few instances of *support*. The following two excerpts contain examples of the face-damaging treatment the men often receive. In 'Lichees' Wu expresses her dismay in line 1 that Ch. has bought both *lichees* and *longan*. Her husband's joking response (line 2) and his repeated attempts to calm her and justify himself (lines 6 and 8) are totally ignored by Wu as she continues to scold him in a loud, fast, wailing manner in lines 3, 5, 7 and 9. In 'Toffee' Sybille's husband, who is already 'in disgrace'⁵⁷ gets himself into more trouble by admitting in line 2 that he has forgotten to buy his daughter some toffee⁵⁸. His excuse in line 2 is

⁵⁵ It is possible that this attitude was in part due to past experiences, which in the case of Wu and her husband was indeed the fact. In the other cases, the experimenter chose not to enquire after this aspect of the couples' relationships.

⁵⁶ See Granet 1958, Baker 1979 and Freedman 1970. A detailed discussion of the position of the Chinese woman throughout the ages is also included in Mueller-Liu 1997.

⁵⁷ This passage starts quite negatively because Cr. was criticized for something else just prior to it.

⁵⁸ It should perhaps be emphasized once more that the interactants themselves determined the course and content of the dialogs. Here, for example, Cr. could just as easily have said that he had bought the toffee. It was his own choice to steer the conversation in this direction, by 'admitting' that he had forgotten it. This shows clearly that even 'only' *quasi-natural* interactions of this kind are the joint product of all the interactants, an important principle in interactional sociolinguistics.

brushed aside by his wife (in line 3) who seems genuinely disappointed at his failure and tells him to explain it to his daughter himself (line 5). His response to this is a little brisk, but has a conciliatory undertone.

'Lichees'

- W: (1) You brought *lichees* and you also bought *longan*?! What on *earth* did you do *that* for?! (very loud, fast, scolding, almost distraught)
- Ch: (2) [For *eating*, of *course*! (smiles) (soft, fast, gentle, joking)
- W: (3) But that's too *much*! *Lichees* and *longan* are the *same kind* of fruit! (as before)
- Ch: (4) [Not *much* (gentle)
- W: (5) [If you bought *lichees* then you shouldn't have bought *longan*, too! (as before)
- Ch: (6) [Not *much* (soothingly, conciliatory) [Oh, no, not *much*, not *much* (as before)
- W: (7) [That must have been really expensive, how *much* were they per kilogram? (as above)
- Ch: (8) [You like them, your *daughter* likes them- (mid-loud, faster, gentle)
- W: (9) [Oh *no*, how *much* were the *lichees* per *kilogram*, you *tell* me! (loud, fast, scolding and wailing)

'Toffee'

- S: (1) And the *toffee*? (mid-fast, a little cool)
- Cr: (2) Oh, you *know*, I didn't bring back any *toffee*. +++ I thought the children + the children already have *so many sweets*
- S: (3) [No *toffee*? (fast, emphatic, disbelieving)
Oh, but Joanne really wanted *toffee so much*. (cool, clearly disappointed)
- Cr: (4) Hm, well (seems at a loss for words)
- S: (5) Well, you can explain that to her yourself. (brisk, matter-of-fact, cool)
- Cr: (6) Okay, I *will*. (equally brisk, but conciliatory)

Sometimes, the women's reaction to a purchase which does not meet with their approval is not to scold their 'husbands' but wail in despair. In most, though by no means in all cases, this 'wailing' strategy has the effect of arousing the husband's sympathy and bringing him round to 'finding' the desired item somewhere among the purchased items. This is shown in the following excerpt, '*Chocolate*'. Here, Clara's wailing complaint in line 3 is first met with a sympathetic 'Oh' from her husband H., who in line 5 then suddenly 'remembers' that he has bought the cooking chocolate after all - and for an excellent price, too. This appeases Clara, who appears happy and relieved in line 8.

'Chocolate'

- C: (1) And the *chocolate*? Did you get the *chocolate*? (a little anxious)
- H: (2) The *chocolate*, of *course*, the *children* asked me to get *that*. I bought - (matter-of-fact)
- C: (3) But I *also* wanted *chocolate* for my *cake*... (mid-fast, louder, wailing)
You *know*, on special offer, for five marks *eighty*...
- H: (4) [Oh (very sympathetic)
- H: (5) Ah *yes*. Of *course*. Actually, I bought that right at the *beginning*. (matter-of-fact, gentle,
And it only cost three marks *eighty*. *Chocolate für icing*. mid-fast, quite in control)
- C: (6) [Mm, Mm]
- C: (7) Oh *good*. (happy, relieved)

In the next excerpt, '*Lemons*', taken from the beginning of the dialog between Wu and her husband, an example is given of the highly disruptive and negative nature of the simultaneous speech found in this interaction. A little confused in lines 1 and 3, Wu suddenly discovers in line 5 that Ch. has bought too many lemons and breaks out in loud lamentations. Ch.'s expression of surprise in line 6 is followed by three attempts to calm her and explain his purchase, which are, however, completely ignored by Wu. After interrupting him in line 7

with three loud exclamations following each other at very brief intervals, she goes on to scold him in such a loud and fast manner that he is quite unable to get a word in edgewise.

'Lemons'

- W: (1) Well, what did you *buy* for me? (fast, mid-loud, confused, seems to be consulting her list)
 Ch: (2) Yes. Well, I bought *everything* you *asked* me to. (mid-loud, fast, smiling, joking)
 W: (3) What- what did I *ask* you to buy? I have no- I' ve already forgotten what
 I *asked* you to buy, let me *see*.. (as before)
 Ch: (4) [Mm. Oh.
 W: (5) Oh, Oh *God*, you bought *lemons*! (fast, very loud, wailing, horrified)
 What did you buy so many *lemons* for, oh *dear!* Oh *dear!*
 Ch: (6) I bought too many *lemons*? [Lem- Lem- Lemons- (soft, mid-slow, gentle, surprised)
 W: (7) [Too many! Too many! Too many!
 W: (8) Lemons, oh dear! Only your daughter wants to drink lemon tea! (as before)

Finally, the excerpt *'Emperor of China'* is an example of the stern and often rather negative treatment You's partner receives. Having discovered that H. has no idea how much the stewed chicken wings cost, You scolds him quite strongly in lines 3, 5 and 7, interrupting him with her loud and fast question in line 5. At first H. seems a little surprised, possibly amused, by the severeness of her tone, laughing a little between lines 4 and after line 5. Then, seemingly warming to the task, he tries to justify himself in line 8 by pointing out that he actually bought a large quantity of stewed meat for the price criticized by You. This, however, does not make things better, as in her opinion stewed meat is too rich and special a dish to be eaten over a long period of time - unless one happens to be the *'Emperor of China'* (line 9).

'Emperor of China'

- Y: (1) Now then, this + stewed *meat* ++ how much was *that*? These stewed *chicken* wings, (mid-slow, mid-loud, matter-of-fact, stern)
 how much did *they* cost?
 H: (2) Stewed *chicken* wings? (slow, hesitating)
 Y: (3) [Mm Mm. How much was this stewed *meat*? (as before)
 H: (4) Hm, I don't *know* (laughs) It -
 Y: (5) [You bought this stewed *meat* and you have no *idea*
 how much it *cost*? (fast, loud, very stern)
 [1.2] (H laughs, Y doesn't)
 H: (6) I think it *cost*- it + stewed *meat* + I think it cost + twenty
marks. (slowly, disjointed, probably consulting his list)
 Y: (7) [You bought] stewed meat for *twenty* marks?! (loud, fast, disbelieving, horrified)
 H: (8) Well, twenty *marks*, it- we can get *five meals* out of it, you *know*. (faster, a little more confident, justifying himself)
 Y: (9) *Five meals!* Are you the emperor of *China*? To eat so *well*?! (mid-loud, scoffing at him, the last part is softer)
 (H laughs again)

To summarize, the interactions between husband and wife are characterized by the strong *footing* and rather negative attitude on the part of the four women. In almost all these interactions, the men are criticized, scolded and generally subjected to extremely face-damaging treatment. The most frequent interaction strategy used by the women here is *arrogance*, with *deference* and *support* being found only occasionally in a few gentler passages of these interactions. *Camaraderie* is virtually absent.

6.2.4 The interactions with unknown men

With the unknown male interaction partners, the women are like completely different people. They are very friendly, polite and respectful, often compliant, at times even demure, giving the impression that in these dialogs they are on their 'very best behavior'. Similar to what was described for the interactions between woman friends, their speech behavior clearly aims to signal goodwill and promote harmony, but here they seem to be making an even greater effort. As in the dialogs with the unknown women the interaction partners must slowly get to know each other in the course of the interaction, but in these dialogs this process does not involve any of the difficulties and critical moments discussed above. With regards to simultaneous

speech, it is interesting to note that while overlaps and interruptions are very frequent in the dialogs between Clara and Li and between You and Hw., they are extremely scarce in the interactions between Sybille and R. and Wu and L. While the reasons for this discrepancy are not quite clear, it is noteworthy that the latter two dialogs are characterized by a slightly stronger feeling of distance and more careful speech behavior on the part of the interactants, which is quite similar to that encountered in the dialogs involving unknown female interaction partners. In contrast to this, the interactions between Clara and Li and between You and Hw. are marked by a very merry and lively atmosphere, which is highly reminiscent of the interactions between woman friends. While in the interaction between You and Hw. the high incidence of interruptions may be explained as part of the teasing, challenging humour strategy of *tai-gang*, which is frequent here, the abundance of interruptions in the dialog involving Clara and Li is more likely to have been due to their unfamiliarity with each other's conversational styles.

As in the dialogs with the woman friends, the interaction strategies of *camaraderie* and *support* are employed very often in these interactions and there is also a very high incidence of *deference*. The enactment of the latter two strategies is particularly striking. In the *deference* strategy all four women use breathy, soft and high-pitched voices to make themselves appear small, helpless and in need of assistance from the men. In the passages containing *support* the offers of assistance are voiced in a very warm, at times almost motherly, manner, the women's voices marked by soft and lax, at times breathy voice quality.

The use of *deference* in these dialogs is illustrated in the following excerpt, 'Goose', where in line 3 Clara's very high pitched and breathy voice signals childlike, almost childish, admiration of Li's goose.

'Goose'

- | | | |
|---------|--|---|
| C: (1) | Now then, now let me s- see what you have in the way of
sausage and meat. | (fast, high, very merry, childlike)
(slows down here and lowers pitch) |
| Li: (2) | [Yes | |
| C: (3) | Ooh, what a beautiful goose! That - | (high, fast, childlike, enthusiastic) |
| Li: (4) | [Oh, the goose, yes | (also enthusiastic) |
| C: (5) | Yes! | (as before) |

The next excerpt, 'Vanilla sugar', shows an example of the *supportive* behavior found in these interactions. Here, partner R suddenly has difficulties (line 3) in pronouncing the word 'vanilla sugar'⁵⁹. Sybille comes to his assistance, first by correcting him in a very warm, motherly way, and then praising him for his efforts (both in line 4).

'Vanilla Sugar'

- | | | |
|--------|---|--------------------------------|
| S: (1) | But I wans't <i>done</i> with the baking things yet... I need vanilla sugar, too. | (soft, mid-slow, warm) |
| R: (2) | [Ah ha. | |
| R: (3) | Ani- er + Anilla sugar, er, Van-, Vanilla sugar. | (slow, disjointed, unsure) |
| S: (4) | [No, Vanilla- [Vanilla sugar, that's right! | (soft, gentle, warm, motherly) |
| R: (6) | Vanilla sugar Yes, we've got that. | |
| S: (7) | [Mm, Mm] | |

Interestingly, despite her extremely demure and *deferent* speech behavior here, the dialog between Wu and her unknown male partner L contains a few passages where the interactants are clearly not on the 'same wavelength'. In the following excerpt, Wu is employing an intensified version of the *deferent* behavior that was discussed above in her interaction with You. Her voice is high pitched and breathy, and she speaks in a very hesitant and disjointed

⁵⁹ This confusion of R., a native speaker of English, who was quite fluent in German, was due to the fact that in the area of Germany where he had lived prior to moving to Saarbrücken the word for vanilla sugar is slightly different (Va' nillezucker).

manner, laughing apologetically every now and then. But, similar to what was found above for her conversation with You, this extremely humble, inferior speech behavior (in lines 1, 4, 5, 7, 9 and 13) clearly does not cut any ice with L and he reacts (in lines 3, 6, 8, 10 and 12) with a cool *distance* reminiscent of You's behavior discussed above in 'Avocados'. All the while L is not only distant but also seems a little unsure what to do or say. While this may have been due to his not knowing how to deal with such extremely submissive behavior from a woman clearly older and more experienced than himself, on the other it may also have been that he was not quite sure himself what *niu-nan* is and did not want to let this show. As a third possibility, it may also have been that L could not tell whether Wu really didn't know what *niu-nan* was or whether she was just putting on a show. In this context Wu's remark in line 11 is of interest: If she did not know what *niu-nan* is, how could she have known that it tastes better than normal beef? Thus, if by 10 partner L had convinced himself that Wu had no idea what *niu-nan* was, after line 11 he was almost certainly unsure again. Indeed, while in lines 8 and 10 he seems relatively confident, in line 12 his speech again becomes disjointed, suggesting that he is again at a loss how to deal with Wu. As a final point, also observed in 'Avocados', L does not join in in Wu's apologetic laughing and smiling (line 13).

'Niu-nan'

- W: (1) Now, this thing called *niu-nan*, what is it? (soft, slow, disjointed, unsure)
 I have *no* idea, what - er, er Yes.
- L: (2) [Oh, *niu-nan*, yes] (mid-loud, matter-of-fact)
- L: (3) It's *beef*- it's- er- it's (disjointed, a little cool? uncertain?)
 a kind of cut of *beef*. It's a kind of *beef*, so to say.
- W: (4) [Beef- a kind of - (mid-loud, a little surer)
- W: (5) Oh, a kind of *beef*...
- L: (6) [Right. (as before)
- W: (7) You mean + it is actually *meat*, not *offal*? (soft, mid-fast, a little unsure again)
- L: (8) Yes, that's *right*. It's *meat*, it's not *offal*. (as before)
- W: (9) Ah, I see ++ Does it taste- er, *similar* to other *kinds* of beef? (as before)
- L: (10) Yes, about the *same*.
- W: (11) About the *same*, meaning it tastes a little *better*... (fast, mid-loud, strangely certain,
 Well *then*, why is it called *niu-nan*? then slower, softer, again unsure)
- L: (12) Er, *niu-nan* + It's *because*- *because*- depending on the different (slow, disjointed, unsure)
cuts of the beef, the *names* are *different*, too.
- W: (13) [Ohh... You *see*, I've really never (soft, sweet, mid-fast)
heard of it before... (laughs apologetically, he doesn't). I'm *so sorry*! (as above)
 Okay, then I'll take a *little* bit + I'll buy + buy a little to *taste* it. (as before, smiles. L still doesn't smile)

Finally, as noted in passing above, in the interaction between You and her unknown male partner Hw. there is a very high incidence of humour, a large proportion of this being the teasing strategy of *tai-gang*. This shows that Chinese *tai-gang* is not restricted to interactions involving friends or interactants familiar with each other, but may also be used by strangers sharing the same conversational style. Again, the underlying message is one of mutual friendship and solidarity. Starting off politely enough, with You on her best *deferent* behavior, the conversation all of a sudden develops into a *tai-gang* match, with both interactants alternately teasing and challenging each other and, judging from the large amount of shared laughter, enjoying themselves very much. The following two excerpts 'Duck meat' and 'Dumplings' are taken from their interaction. After having suggested that Hw. buy some duck meat (in line 1) and having praised its good quality (line 3), You tries the same ploy as with F above: After stating in line 5 that duck meat is in season and therefore cheap, in the same breath she gives Hw. a price that is anything but cheap. This is delivered in an emphatic and challenging way which Hw. picks up very quickly, returning the challenge immediately, in line 6. After they have squabbled to and fro for a few minutes (in lines 7 - 9), Hw. tries a new ploy, informing You (in 10) that he will wait with his purchase of duck meat until it has become less expensive. The price he names, eight Marks, is however so low that You cannot help but

make fun of him in line 11 by telling him that by that time, the duck will have become rotten and inedible.

'Duck meat'

- Y: (1) Well, what about some *duck* meat or *goose* meat? (mid-fast, mid-loud, merry, friendly)
Hw: (2) Er + *duck*...meat...Is this *fresh* duck meat or (mid-soft, slow, disjointed, unsure)
is it deep-frozen?
Y: (3) [Oh, *fresh*, very *fresh*. We don't sell *frozen* duck meat. (loud, fast, enthusiastic)
Er, our ducks are always *fresh*, still *warm* in *fact*! (both laugh)
Hw: (4) Okay, er, how much does it *cost* at the *moment* ? (smiling, mid-fast)
Y: (5) At the *moment*, duck is in *season*, very *cheap*. Twenty marks (fast, loud, enthusiastic)
per *kilogram*!
Hw: (6) Oh, that's *cheap*, is it?! (smiles) (fast, emphatic, challenging)
Y: (7) [Oh *yes*! (Smiling) (as above)
Hw: (8) [That's really *expensive*! (Smiling) (mid-loud, mid-fast, jokingly protesting)
Y: (9) Well, if it were *out* of season, it would cost even *more*! (both laugh) (as above)
H: (10) Okay, then, I'll - I guess I'll *wait* a few months and buy it (mid-loud, mid-fast, smiling)
when it's less *expensive*. I'll wait until + it
costs *eight* marks per *kilo*.
Y: (11) [Oh, I see. Well, in *that* case, I'm afraid you may (as before)
have to wait until the the *whole place* smells of [*rotten*] *duck*! (both explode with laughter)

In the following excerpt 'Dumplings, the procedure is similar. Hw. begins the passage by stating that he does not need to buy dumplings, as he can make these himself. Since this is an indirect attack on You's positive face⁶⁰, in line 3 she immediately tries to persuade him into buying some after all by calling attention to their good quality and their famous red bean paste filling. Hw. seems a little confused at first (line 4), but then 'attacks' in line 6 by pointing out to You - the 'vendor' - that since they have this filling, they cannot possibly be *man-tou* but must be called *bao-zi*⁶¹. Given the fact that she is the 'vendor' and may be assumed to know what she is selling, this amounts to a very strong attack on her positive face, as Hw. is thus insinuating that she doesn't know what she's talking about. You, however, is unperturbed, she simply denies this (line 7) and insists on calling her merchandise '*man-tou* filled with red bean paste'. As she is now suggesting that it is Hw. who has no idea what the difference between *bao-zi* and *man-tou* is, this is quite as strong an attack on his positive face as his observation in line 6 was on hers.

'Dumplings'

- Hw: (1) *Man-tou* dumplings I can make *myself*. I don't need *man-tou* dumplings. (mid-slow, mid-soft, pensive, smiling)
Y: (2) Hm
Y: (3) Oh, I see. You don't want to *taste* ours? They're *very famous*! Like, (mid-fast, mid-loud, challenging)
these with red *bean* filling- red *bean* paste, they're *delicious* + What about
trying them?
Hw: (4) Er, with red *bean* paste (slow, mid-soft, a little disjointed, unsure)
Y: (5) [Right! (as before)
Hw: (6) Those- those are called *bao-zi*, not *man-tou* ! (laughs) (fast, mid-loud, challenging)
Y: (7) Oh, no, absolutely *not*,
[They're called] *man-tou* filled with red *bean* paste! (both laugh) (fast, emphatic, challenging)

To conclude, in the encounters between the speakers and the male partners unknown to them, the speech behaviour of the speakers aims at communicating friendship and goodwill and generally alternates between *deference* and *support*, with *camaraderie* also being present. While these strategies were also employed in the interactions with the known and unknown women, it does seem that here the women invest a greater effort, often giving rise to seemingly exaggerated deferent and supportive behavior. In the interaction between You and her unknown male partner Hw. there is a high incidence of *tai-gang* which, though signaling

⁶⁰ This refusal of You's *man-tou* amounts to an attack on her positive face as Hw. is thus expressing the wish for dumplings other than hers.

⁶¹ *Bao-zi* and *man-tou* are similar in consistency and taste, but *bao-zi* are usually smaller and pointed, *man-tou* larger and more like buns.

While in Taiwan both *man-tou* and *bao-zi* can have this filling, in Mainland China this is not the case.

(playful) *arrogance* on the surface, is meant to communicate *camaraderie* - and is also understood as such⁶².

6.2.5 Summary and conclusions

The above discussion of the recorded interactions has yielded the following findings: First, despite the fixed outer frame of the interactions and the predetermined topic-matter, the participants in the dialogs were clearly able to act and interact with each other naturally. Second, true to their assignment, the interactants remained completely themselves during the dialogs, mapping their own personalities and conversational styles onto their roles of 'vendor', 'customer' and 'wife'. The feelings, reactions and attitudes which were built into the dialogs were genuine and not simulated in any way. Third, the attitudes of the speakers varied between the dialogs, a fact which is reflected in their negotiation of the interactional events which were enacted in each dialog. And fourth, while this negotiation sometimes also varied slightly in the course of one dialog, it is possible to broadly describe each of the four types of interaction in terms of one general underlying feeling or atmosphere and in terms of the interaction strategies most often employed therein. Notwithstanding minor differences across the two languages, the most important of this involving the use of *tai-gang* humour, found only in the dialogs involving the Taiwanese speaker You, each of the four settings is negotiated in much the same manner by the German and the Chinese speakers. Table 1 below lists the speakers' roles, the identity of their interaction partners and the duration of the dialogs in each of the German and Chinese interactions.

<i>Speaker</i>	<i>Setting</i>	<i>Role of speaker</i>	<i>Interaction partner</i>	<i>Duration</i>
Clara	unknown man	'customer'	'Li'	8 mins.
	familiar man	'wife'	Husband 'H'	9 mins. 30 secs.
	unknown woman	'vendor'	'Wi.'	13 mins.
	familiar woman	'customer'	Sybille	14 mins. 30 secs.
Sybille	unknown man	'customer'	'R.'	12 mins.
	familiar man	'wife'	Husband 'Cr.'	8 mins.
	unknown woman	'vendor'	'Wi.'	9 mins. 12 secs.
	familiar woman	'vendor'	Clara	14 mins. 30 secs.
You	unknown man	'vendor'	'Hw.' (China)	9 mins.
	familiar man	'wife'	Boyfriend 'H.'	9 mins.
	unknown woman	'customer'	Wu (China)	6 mins. 10 secs.
	familiar woman	'vendor'	'F.' (Taiwan)	9 mins. 5 secs.
Wu	unknown man	'customer'	'L' (Taiwan)	6 mins. 30 secs.
	familiar man	'wife'	Husband 'Ch.'	7 mins. 35 secs.
	unknown woman	'vendor'	You (Taiwan)	6 mins.
	familiar woman	'customer'	S. (China)	6 mins. 5 secs.

Table 1 Speaker roles, interaction partners and duration of the recorded dialogs

⁶² These examples, which may be regarded as examples for relative 'tame' *tai-gang*, show that in this strategy the intended metamessage is, as Tannen also notes of 'ritual complaining', in fact one of *camaraderie*. According to Norrick (1993:44), this does not apply to the English use of irony and sarcasm in joking. As another difference to English, also evidenced here, *tai-gang* in China may be used not only with familiar interaction partners, but also with total strangers.

The comparative analysis of the German and Chinese interactions has yielded the following findings: In both languages the merry, harmonious interactions between woman friends are characterized by the employment of genuine *deference*, warm, hearty *support* as well as a relatively high incidence of humour accompanied by shared laughter. The less harmonious interactions between the women unknown to each other seem to be characterized by a careful, not always entirely friendly distance. Though *deference* and *supportiveness* are also used here, in a number of cases their undertones seem to lack warmth and sincerity. *Camaraderie* is rare and a number of utterances by one or the other of the interactants have a distinctly cool and superior feeling, suggesting the presence of *arrogance* here. In both the German and the Chinese interactions between the speakers and their husbands/boyfriends, the attitude of the speakers is one of suspicion mixed with cool, often very negative undertones. *Camaraderie*, *support* and *deference* are rare, and in some of these interactions absent. Occupying a very superior *footing*, the women all question, criticize and often scold the men in a highly insensitive manner which, to all appearances, quite genuinely signals *arrogance*.

The very few exceptions involve the *tai-gang* strategy of humour, evidence of which is strong in the Taiwanese interaction, between You and her boyfriend. Finally, in stark contrast to this, the speech behavior of the German and Chinese women towards the men unknown to them is extremely warm, friendly and considerate, in fact friendlier and more polite than with their women friends. Clearly intending to make the interactions as pleasant as possible for the men, the four speakers use *deference* and *support* strategies whose overall feeling - respectful, demure, at times childlike on the one hand, warm and motherly on the other - often seems overly sincere and thus slightly exaggerated. As a special feature, there is a high incidence of *tai-gang* humour in the dialog between You and her unfamiliar male partner. These findings indicate that the German and Chinese women display very similar, at times identical, interactional behavior in all four types of dialogs. On the basis of the above findings it is concluded that the interaction strategies of *camaraderie*, *deference*, *support* and *distance*, discussed in Chapter 1 for English interpersonal interaction, are equally known in Germany and China and are used there in much the same manner.

Finally, the number of interruptions and simultaneous speech was briefly examined and found to vary strongly between the interactions. Broadly speaking, the results indicate that the cooler dialogs are marked by a lower incidence of simultaneous speech, while the warmer, merrier interactions, especially those involving female interaction partners, have a relatively high number of simultaneous speech and interruptions. Many interruptions were also found in the interaction between You and her unfamiliar male interaction partner, also marked by a high incidence of the challenging humour strategy *tai-gang*. Table 2 shows the number of interruptions and instances of simultaneous speech found in each interaction.

The above informal analysis of the German and Chinese recorded dialogs centered on the most frequent interactional events and the characteristic ways in which these were negotiated by the four speakers, this in turn serving as a rough indication of the interaction strategies used in each case. A more objective account of the speakers' speech behavior was sought in the sociolinguistic judgement tests, to be discussed in the next chapter, by asking native speakers of German and Chinese to analyse certain utterances taken from the dialogs in terms of the interaction strategy signaled thereby⁶³.

⁶³ The following 10 speech samples to be discussed in Chapter 7 originated in the passages discussed above: *cSchoko 2* ('Chocolate'), *cGans 1* ('Goose'), *cZwieb 3* ('Summer Onions'), *cMango 3* ('Mangos'), *sMango 4* ('No Mangos'), *yLizhi 4* ('Eight Marks'), *yMan 1* ('Dumplings'), *wNing 2* ('Lichees'), *wNing 2* ('Lemons') and *wLa 4* ('Inviting Guests').

Type of setting	Interaction	Interruptions	Simult. speech	Total
Familiar woman	Clara & Sybille	8	21	29
	Wu & S.	14	9	23
	You & F.	8	8	16
Unknown woman	Clara & Wi.	11	12	23
	Sybille & Wi.	4	5	9
	Wu & You	4	9	13
Familiar man	Clara & H	4	8	12
	Sybille & Cr.	5	5	10
	You & H.	12	11	23
	Wu & Ch.	9	7	16
Unknown man	Clara & Li	8	20	28
	Sybille & R.	2	6	8
	You & Hw.	6	11	17
	Wu & L	1	4	5

Table 2 Interruptions and simultaneous speech in the recorded dialogs

7. The sociolinguistic judgement tests

The preceding macroanalysis of the recorded dialogs was based in the main on the personal observations of the experimenter and her familiarity with the personalities and conversational styles of the participants, in particular the four speakers¹. Useful as this knowledge was for the examination of the naturalness of the speaker's speech behavior and a first, tentative analysis of the interactional events in terms of the attitude signaled by the speakers, it was, however, unsuitable for the determination of the interaction strategy signaled by individual utterances, as conceived by someone unacquainted or less familiar with the speakers². This objective analysis and classification of selected utterances in terms of speaker attitude, i.e. interaction strategy, was made the object of the judgement tests to be discussed in this chapter. In the following we will first discuss the general procedure of the tests, including the participants, the speech material and the questions, before coming to the evaluation of the tests, leading to the classification of the selected utterances, or speech samples, in terms of type of interaction strategy. The third point to be discussed will be the influence of the discourse context on judgements of the utterances.

7.1 The general procedure of the judgement tests

7.1.1 The participants

A total of 59 judges, all native speakers of German or Chinese between 25 and 35 years old, were asked to participate in the judgement tests. Most were selected by the experimenter, but in a few cases judges were also introduced by speakers whom they were acquainted with. Most of the German participants were from the Saarland or the Rhineland-Palatinate area, and all were familiar with the Saarland variant of High German. The Chinese and Taiwanese judges, all either studying or working in Saarbrücken, had been living in Germany for an average of two years. While they all spoke German moderately well, since they had very little private contact with Germans and rarely spoke German in their free time, it may be assumed that their Chinese was still fluent and also remained unaffected by phonetic/phonological influences from the German language. With very few exceptions, most of the participants in these tests were linguistically naive and none were informed about the actual goal of the tests.

7.1.2 The speech material

The speech material consisted of 103 German and Chinese speech samples, each consisting of one, maximally two utterances produced by one of the four speakers³. Each speech sample contained a *keyword*, corresponding to one of the articles on the shopping-list used in the recorded dialogs, and used in the labeling of the speech samples, to be described below. The utterances were all lexically unmarked with respect to the attitude expressed therein, i.e. they contained no lexical factors, such as hedging devices or forceful expressions, which might have 'given away' the attitude intended by the speaker at the time.

The speech samples had been selected from the recordings by the experimenter after a very thorough auditory analysis⁴, for one of the following two reasons: According to her judgement, each either expressed very clearly one of the interaction strategies support, arrogance, deference or false deference, or contained strong ambiguities as to speaker attitude in terms of these interaction strategies. All the speech samples had been transferred onto audiotapes, once

¹ As was seen in the preceding chapter, in the macroanalysis the participants' own views of interactional events were solicited whenever this was necessary - and possible - for the clarification of certain passages.

² As has been pointed a number of times, in the judgement tests it was not a matter of determining exactly what the speaker had intended with a given utterance, but rather how it would be judged objectively by listeners on the basis of its suprasegmental features.

³ Of these utterances, 24 had been spoken by Clara, 25 by Sybille, 30 by You and 24 by Wu.

⁴ Among the criteria applied in the selection of these utterances were: (1) No slips of the tongue (2) No excessively long stretches of silence (3) No simultaneous speech and (4) a 'neutral' content, i.e. containing lexical or syntactic indications of the intended interaction strategy.

with and once without their preceding discourse context, in all cases, however, excluding the responses of the interaction partners. This was meant to ascertain that the judges would have no knowledge of how he/she had originally interpreted the speaker's utterances and would be unbiased in their own judgements of speaker attitude.

7.1.3 *The procedure of the tests*

The 59 German and Chinese judges were divided into 4 groups, or panels, of 14 or 15 members, each group being assigned the utterances of one of the four speakers. Each panel of judges consisted to approximately equal parts of men and women and judges familiar and unfamiliar with the speaker. The tests were held separately for each panel in small groups of three or four judges at a time, either at the phonetics institute or the home of the experimenter. The judges' task was to analyse every speech sample with the help of a questionnaire, first in broad terms, then in more detail, focusing on the attitude - the interaction strategy - which in their opinion the speaker was signaling in each case, the choice being between *support*, *arrogance*, *deference* and *false deference*⁵. Before the beginning of each test, the questionnaires were distributed, read aloud and discussed in the language of the judges. It was emphasized that their judgements were to be based not so much on the lexical and grammatical content of the utterances, but as much as possible on the effect of their suprasegmental features, this being explained as relating to the general 'feeling' of the utterances, i.e. 'not what the conversationalists said but how they said it'. After this, all the speech samples of the respective speaker were played to the judges once, in a block, so that they could gain a general impression of the speaker's voice as well as a clearer understanding of what they were being asked to do. The test itself which then commenced consisted of two parts, separated by a short interval. In the first half of the test the speech samples were offered in randomized order and in isolation, while in the second half, the utterances were played to the judges together with their respective discourse contexts consisting of three to five conversational exchanges immediately preceding the speaker's utterance. In the first part of the test each utterance was played to the judges repeatedly, the average being about six times, in the second half this number was reduced to twice or three times.

During the approximately half-hour interval which separated the two parts of the judgement test, refreshments were served and everyone was asked to relax, chat and put the test out of their minds. Discussions relating to the test were avoided during this time, so that everyone would forget the responses they had given and proceed to the second half of the test in an unbiased manner. On the other hand, this change of frame was also meant to loosen the atmosphere, lift the pressure on the participants and increase their motivation. After the completion of the judgement test, the participants were invited to voice their impressions and comments with respect to either particular utterances or certain aspects of the test itself⁶.

7.1.4 *The questionnaires*⁷

The questionnaires, designed to guide the judges in their analysis of the speech samples, were written in their respective native language, i.e. in Chinese or German, and contained four multiple choice questions for each utterance which the judges were to answer in a pre-determined order. The first two questions, of a more general nature, were meant to slowly hone the judges to their task. Thus, question one asked about the general atmosphere or

⁵ As it was impossible to explain these concepts to the judges in detail in such a short time, the questionnaire had been designed to guide them in the accomplishment of the tasks.

⁶ This included, most importantly, the naturalness of the utterances and the speech behavior of the speakers. As a point to be remembered for later on, many participants called attention to the fact that they had often judged utterances with and without context quite differently. As a second point, to which we will return in Chapter 9, a number of the Chinese judges also commented on one particular utterance spoken by Wu, in which she displayed what seemed to them a rather petulant and childish behavior.

⁷ The questionnaire used in these judgement tests is shown in Section 1 of Appendix II (English translation).

feeling of the utterance, including the presence of humour or irony, such as for example *tai-gang* in the Chinese speech samples, while in the second question, the judges were asked to voice their opinion about the speaker's general attitude towards his interaction partner, the choice here being between cordiality, friendliness, warm or cool distance, coolness and coldness. After having completed the first two questions, the judges were then asked to proceed to questions 3 and 4, designed to assist them in the analysis of *footing* and *valence*, the two components of the metamessage. Question 3 asked about what was referred to as the 'footing', or 'power position', of the speaker with respect to the listener, the choice being between very superior/dominant, superior, equal position, inferior/weak and very inferior/submissive. And in Question 4 the judges were asked to describe what was termed the 'precise' attitude of the speaker towards the interaction partner as a function of two feature-pairs: (1) positive and negative undertone, corresponding to positive and negative *valence*, and (2) strength of the speaker's position, explained as corresponding to the strength, or force, with which the utterance appeared to have been delivered⁸. Strength of position was divided into 3 degrees, weak, strong and very strong, among which the judges were to choose in their responses. While the judges' responses as to undertone corresponded to the *valence* of the speech samples, two responses were used for the determination of *footing*: those relating to 'power position' and those relating to 'strength of position'⁹. In principle, the questions were forced choice questions, meaning that each question was to receive one - but only one - response and that no 'neutral' options were offered¹⁰. However, the judges were permitted to leave out questions they found too difficult, and when answering Question 4, they were also allowed to give more than one response if they felt they really could not decide¹¹.

7.2 The evaluation of the judgement tests - the quantitative analysis

After the tests, the judges' responses were examined by the experimenter, concentrating on the responses given for questions 3 and 4. Speech samples which the judges had found to sound unnatural as well as those which had contradictory, unclear or no responses for these two questions were withdrawn from the speech corpus, this selection yielding a total of 74 utterances whose results were evaluable¹². After this, the results for Questions 3 and 4 in the remaining speech samples were subjected to a quantitative analysis with the following three objectives: First, based on the judges' responses, each speech sample was to be classified in terms of one, maximally two, values, or scores, of *footing*/strength of position and *valence*, combinations of these scores making it possible to assign each speech sample to one, maximally two, types of interaction strategy; Second, the consistency of the judges' responses, reflected in the standard deviation for each speech sample and seen as an index of its possible ambiguity, was to be determined¹³; And third, systematic differences between judgements made with and without the discourse context were to be sought.

⁸ What is referred to here as 'strength of position' corresponds to what was termed in Chapter 2 the manner in which a message is delivered by the speaker. In the semantic differential model of Osgood et al. (1957) it corresponds to *activity*, the amount, or strength, of feeling or interest of a speaker, either towards his message or his interaction partner.

⁹ The responses to both *footing* and strength of position were taken as the basis for *footing* to even out possible discrepancies between these responses, reflecting fluctuations in the judges' opinion, and thus increase the degree of objectivity of the judgements. As another reason, it was thought that this would also reduce the number of 'neutral' responses for *footing*, which proved to be correct.

¹⁰ The application of 'forced choice' and the relative absence of 'neutral' options was inevitable: In the evaluation of pilot listening tests, which contained 'neutral' options and thus did not apply the 'forced-choice' principle, it was discovered that a large number of the judges, especially the Chinese ones, had - possibly out of consideration for the speaker's face - chosen the 'neutral' option for virtually all the questions in all but a very small number of the speech samples. As this abundance of 'neutral' responses resulted in equal *footing*, unspecified *valence* and 'mid' strength of position for almost all the speech samples of the Chinese speakers, responses of this kind were not useful for the experiment. 'Forced choice' questions were subsequently introduced to avoid their occurrence as much as possible.

¹¹ Ambiguity in Question 4 was in fact not so serious, as this could serve as an indication of the potential ambiguity of speech samples.

¹² Of these, 20 speech samples had been uttered by Clara, 17 by Sybille, 18 by Wu and 19 had been produced by You.

¹³ While a low *standard deviation* is indicative of a relatively clear and uncontroversial classification of the corresponding utterance(s) in terms of *interaction strategy*, a high *standard deviation*, reflecting (strong) inconsistencies in the judges' responses, gives reason to assume (a certain degree of) ambiguity in the utterance.

For the quantitative analyses, the judges' responses for *footing*, strength of position and *valence* were assigned values, or scores. The responses for *footing* received the following scores: '1' and '2', corresponding to *very submissive* and *submissive* footing, '4' and '5', corresponding to *superior* and *very superior*, or *dominant* footing and '3', corresponding to *equal footing* between interactants. The scores chosen for *valence* were '1' for *positive* and '0' for *negative* valence. In the case of strength of position, '1' denotes *weak*, '2' *strong*, and '3' *very strong* strength of position.

In the following, we will first discuss the results of the classification of the speech samples according to type of interaction strategy, carried out on the basis of judgements made without the discourse context. As the potential ambiguity of the speech samples is of interest, brief mention will also be made of values for standard deviation and the consistency of the judges' responses. As a second point, we will discuss the results for the influence of the discourse context on the judgements.

7.2.1 The classification of the speech samples¹⁴

7.2.1.1 Procedure

Based on the system laid out above, the judges' responses to questions 3 and 4 for each speech sample were first given scores relating to *footing* (1 - 5), *valence* (0 or 1) and strength of position (1 - 3)¹⁵. After this, the mean scores of these responses were calculated, every utterance receiving a mean score for *footing*, one for *valence* and one for strength of position. As a third step, based on these mean scores, the speech samples were then classified in terms of type of interaction strategy. While the numerical representation was retained for *footing* and strength of position, mean scores for *valence* were converted into positive and negative mathematical signs: Scores >0.5 were converted into a plus sign (+), denoting positive *valence*, while scores <0.5 were converted into a minus sign (-), representing negative *valence*. To give two examples of this classification, mean scores of 3.6 for *footing* (superior), 2.6 for strength of position (rather strong) and 0.2 for *valence* (negative), yielded a final overall score of [3.6;-2.6], corresponding to superior *footing*/strength of position and negative *valence*, seen as representing *arrogance*. An overall score of 2.0 for *footing* (inferior), 1.2 for strength of position (weak) and 0.8 for *valence* (positive) yielded an overall score of [2.0; +1.2], corresponding to inferior *footing*/strength of position and positive *valence*, taken to represent the strategy of *deference*. Strong inconsistency in the judges' responses with respect to *valence*, indicating potential ambiguity and reflected in a mean score of around 0.5, was represented with the symbol '+/-'. While ambiguities with respect to *footing*/strength of position are not visible in the mean scores for these categories, they are reflected in the standard deviations of the utterances, mean scores of ≥ 0.9 indicating potential ambiguity. The results of these calculations are listed in Tables A - D of Annex II, Section 2. The row named 'x~' shows the mean scores for *footing* (in column 1), for strength of position (column 3a) and *valence* (column 4a) for each utterance, while the row denoted 'sd' shows the standard deviation of the responses for each of these categories, in the same columns¹⁶.

7.2.1.2 The mean scores for *footing*, *valence* and strength of position

Tables 3 and 4 below list the mean scores for *footing*, *valence* and strength of position for the 74 German and Chinese speech samples. The labels of the utterances consist of their *keyword*, or an abbreviation thereof, followed by a number designating the setting in which it was

¹⁴ The first results to be discussed here thus to judgements made without knowledge of the discourse context.

¹⁵ In cases where judges had been unable to decide and therefore given two responses, the mean value of the indicated responses was taken.

¹⁶ This was performed with an SPSS package.

spoken¹⁷. While 1 and 3 refer to interactions with the unfamiliar male and unknown female interaction partner, 2 and 4 designate the dialogs with the familiar man and woman, respectively. The letters *c*, *s*, *y* and *w* at the beginning of the labels relate to the name of the speaker.

*The scores for footing*¹⁸

The examination of the responses for *footing* (question 3) revealed that while the judges had been able to analyse most of the utterances without problems, for each speaker there were a number of speech samples whose results were a little inconsistent, with standard deviations of ≥ 0.9 , indicating potential ambiguity with respect to *footing*. In Table 3, which lists the calculated mean scores for *footing* of all the German and Chinese utterances, ordered according to their relative strength, these (26) potentially ambiguous utterances are marked with the symbol ‘*’. In Table 3 scores below 2.0 are referred to as ‘very submissive’, scores between 2.0 and 3.0 as ‘submissive’, scores for *footing* between 3.0 and 3.5 as ‘neutral - a little dominant’ and scores above 3.5 as ‘dominant’. The standard deviations of the potentially ambiguous speech samples appear below the scores for each speaker.

The scores for valence and strength of position

The examination of the judges’ responses for *valence* and strength of position yielded similar results as for *footing*. As Table 4 shows, responses were inconsistent with respect to *valence* in 16 cases, while in 34 cases slight inconsistencies were found for strength of position. In Table 4 these speech samples are all marked with the symbol ‘*’, potential ambiguity with respect to *valence* being marked with the symbol ‘+/-’. The standard deviations for strength of position, in almost all cases relatively low, ≤ 0.7 , are not included here¹⁹, but the distribution of the judges’ responses for *valence*, indicative of various degrees of ambiguity with respect to this category is indicated between the tabular sections. In Table 4, the speech samples are ordered according to their relative strength and their positive or negative *valence*. Scores for strength of position ≤ 1.5 are referred to as ‘weak’, scores between 1.6 and 2.0 are ‘fairly strong’, and scores above 2.0 are taken as ‘very strong’.

7.2.1.3 The ambiguity of the speech samples

As Tables 3 and 4 show, ambiguous results for *footing*, *valence* and strength of position were found in a number of cases, i.e. 26 for *footing* and 16 for *valence* and 34 for strength of position. In order to gain a clearer understanding of the degree of ambiguity present in each case, the response distributions for each individual judgement were examined closely. This revealed that truly strong inconsistencies are present in only a few cases. Thus, with respect to *footing*, among the 26 speech samples with standard deviations of ≥ 0.9 , response distribution is relatively clear-cut (i.e. either 9: 7 or 10: 6) in all of Clara’s 20 speech samples, and among Sybille’s 17 speech samples only six - *sGans 1*, *sBlum 1*, *sSchoko 1*, *sKiwi 1*, *sBlum 2* and *sKiwi 2* - show less clear-cut tendencies among the judges’ responses (8:7). Finally, among Wu’s 18 utterances, all but 4 have clear-cut response distributions, strongly divergent responses being apparent only in *wDan 12* and *wDan 11*, with response distributions of 7: 7, and in *wLa 4* and *wBing 3*, where the judges’ opinions diverge strongly, at a ratio of 3: 3: 3²⁰. Finally, of You’s 19 speech samples only *yMan 1* has a response distribution of 7: 7. This means that a total of 11 German and Chinese speech samples may be regarded as containing strongly inconsistent responses and thus as potentially ambiguous with respect to *footing*: *sGans 1*, *sBlum 1*, *sSchoko 1*, *sKiwi 1*, *sBlum 2*, *sKiwi 2*, *wDan 12*, *wDan 11*, *wLa 4*, *wBing 3*

¹⁷ The labels *wDan 11* and *wDan 12* denote two speech samples containing the same keyword and originating in the same setting (1).

¹⁸ The mean scores of utterances for *footing*/strength of position and valence will henceforth be simply referred to as scores.

¹⁹ These may be found in Tables A - D of Section 2, Appendix II, to be discussed below.

²⁰ Among Wu’s speech samples, *wLu 4* and *wLizhi 3* have response distributions of 8: 6.

and *yMan 1*²¹. With respect to strength of position, responses for this feature were also found to be relatively consistent in most cases and standard deviation is ≤ 0.7 in all but three utterances. Though standard deviation is a little higher, i.e. 0.8, in the utterances *yBing 4* and *yLong 4*, their responses for strength of position do show clear tendencies, 10: 5 and 9: 6, respectively. Only one of Wu's speech samples, *wNing 4*, with a standard deviation of 0.9 for strength of position and a response distribution of 7:7, may be considered as truly potentially ambiguous with respect to strength of position. Finally, 16 utterances were found to have high standard deviations, ≥ 0.9 , for *valence*. Based on the distribution of their responses, indicated in Table 4, it may be said that *cKuli 2*, *cZwieb 2*, *sBlum 1*, *sSchoko 3*, *sBlum 2*, *sSchoko 4*, *wNing 2*, *wDan 11*, *wLu 4* and *wMan 3* may be regarded as as highly ambiguous, while *cSchoko 2*, *cBlut 3*, *cVani 2*, *sGans 1* and *sMoon 2* and *yMing 1*, with response distributions of 8: 7 or 9: 6, are slightly ambiguous for *valence*. Given the important role of *valence* in the differentiation between *support* and *arrogance*, and *deference* and *false deference*, all of these speech samples were taken as potentially ambiguous for *valence*.

To summarize, a first evaluation of the results of the tests based on values for standard deviation had led to the discovery that potential ambiguity with respect to *footing*, strength of position and *valence* is found in a total of 76 judgements. A closer scrutiny of the results, however, in terms of the *response distributions* for each individual judgement, revealed that the large majority of response distributions are in fact relatively clear-cut in so far as most responses point in one direction, i.e. either strong or weak (for *footing* and strength of position) and either positive or negative (for judgements of *valence*). Based on these response distributions, strong evidence for ambiguity with respect to strength of position was found only once, in *wNing 4*, while marked inconsistencies in the responses for *footing* were found in eleven speech samples (*sGans 1*, *sBlum 1*, *sSchoko 1*, *sKiwi 1*, *sBlum 2*, *sKiwi 2*, *wDan 12*, *wDan 11*, *wLa 4*, *wBing 3* and *yMan 1*) and strong indications for ambiguity with respect to *valence* was found in 16 utterances (*cKuli 2*, *cZwieb 2*, *sBlum 1*, *sSchoko 3*, *sBlum 2*, *sSchoko 4*, *wNing 2*, *wDan 11*, *wLu 4*, *wMan 3*, *cSchoko 2*, *cBlut 3*, *cVani 2*, *sGans 1*, *sMoon 2* and *yMing 1*). A total of 25 speech samples were affected by these discrepancies.

²¹ In all these cases, response distributions are either completely mixed or 8: 8.

Clara:	very submissive	submissive	neutral - a little dominant	dominant
	<i>cBlum 4</i> (1.8)	<i>cSchoko 2</i> (2.3)	<i>cMoon 3</i> (3.2)	<i>cSchoko 3</i> (3.7)
	<i>cZwieb 3</i> (1.7*)	<i>cKuli 1</i> (2.3)	<i>cMango 3</i> (3.1*)	<i>cKuli 3</i> (3.8)
		<i>cMoon 1</i> (2.5*)	<i>cMango 2</i> (3.1)	<i>cBlum 1</i> (3.8)
		<i>cGans 1</i> (2.6)	<i>cBlum 3</i> (3.3*)	<i>cVani 2</i> (3.9)
		<i>cKuli 2</i> (2.8*)	<i>cBlut 3</i> (3.5*)	<i>cMango 1</i> (4.1*)
		<i>cGans 4</i> (2.9)	<i>cBlut 4</i> (3.5*)	<i>cZwie 2</i> (4.0)
	Standard deviations:	<i>cZwieb 3</i> (0.9); <i>cMoon 1</i> (1.0); <i>cKuli 2</i> (0.9); <i>cMango 3</i> (1.1); <i>cBlum 3</i> (0.9); <i>cBlut 3</i> (1.1); <i>cMango 1</i> (0.9); <i>cBlut 4</i> (1.1)		
Sybill:	very submissive	submissive	neutral - a little dominant	dominant
	None	<i>sKuli 1</i> (2.4*)	<i>sBlum 1</i> (3.1*)	<i>sMoon 2</i> (3.6)
		<i>sMango 4</i> (2.8)	<i>sKiwi 4</i> (3.1)	<i>sKiwi 2</i> (3.6*)
		<i>sGans 1</i> (2.9*)	<i>sSchoko 3</i> (3.4)	<i>sKiwi 3</i> (3.9)
			<i>sSchoko 1</i> (3.4*)	<i>sSchoko 4</i> (3.6)
			<i>sMoon 1</i> (3.4)	<i>sKuli 4</i> (3.8)
			<i>sBlum 3</i> (3.5)	<i>sMango 3</i> (4.4)
			<i>sKiwi 1</i> (3.5*)	
			<i>sBlum 2</i> (3.3*)	
	Standard deviations:	<i>sKuli 1</i> (0.9); <i>sGans 1</i> (0.9); <i>sBlum 1</i> (1.1); <i>sSchoko 1</i> (0.9); <i>sKiwi 1</i> (0.9); <i>sBlum 2</i> (1.0); <i>sKiwi 2</i> (0.9)		
Wu:	very submissive	submissive	neutral - a little dominant	dominant
	<i>wNing 4</i> (1.8)	<i>wLizhi 4</i> (2.3)	<i>wLizhi 1</i> (3.0)	<i>wMan 4</i> (3.7*)
		<i>wNing 2</i> (2.4*)	<i>wBing 3</i> (3.0)	<i>wMan 2</i> (3.8*)
		<i>wLu 4</i> (2.6*)	<i>wBing 1</i> (3.0)	<i>wBing 2</i> (4.5)
		<i>wLa 4</i> (2.7)	<i>wDan 12</i> (3.4)	<i>wLizhi 2</i> (4.8)
		<i>wMan 3</i> (2.8*)	<i>wDan 11</i> (3.2)	
		<i>wLizhi 3</i> (2.8*)	<i>wLa 3</i> (3.3)	
			<i>wDan 4</i> (3.3)	
	Standard deviations:	<i>wNing 2</i> (1.5); <i>wLu 4</i> (0.9); <i>wLa 4</i> (1.1); <i>wMan 3</i> (0.9); <i>wLizhi 3</i> (1.0); <i>wBing 3</i> (1.3); <i>wDan 12</i> (1.1); <i>wDan 11</i> (0.9)		
You:	very submissive	submissive	neutral - a little dominant	dominant
	None	<i>yLong 3</i> (2.4)	<i>yMing 3</i> (3.1)	<i>yLu 2</i> (3.6)
		<i>yLu 3</i> (2.7)	<i>yMan 3</i> (3.1)	<i>yLong 1</i> (3.6*)
		<i>yLizhi 4</i> (2.7)	<i>yLing 1</i> (3.2)	<i>yMan 2</i> (4.1)
		<i>yBing 4</i> (2.9*)	<i>yLizhi 1</i> (3.3)	<i>yLing 2</i> (3.6*)
			<i>yLong 4</i> (3.3*)	<i>yLong 2</i> (4.3)
			<i>yMing 1</i> (3.3*)	<i>yLa 2</i> (4.4)
			<i>yMan 1</i> (3.3)	<i>yLizhi 2</i> (4.6)
			<i>yDang 1</i> (3.3)	
	Standard deviations:	<i>yBing 4</i> (0.9); <i>yLong 4</i> (1.1); <i>yMing 1</i> (0.9); <i>yMan 1</i> (0.9); <i>yLing 2</i> (0.9); <i>yLong 1</i> (1.0)		

Table 3 The results of judgement tests: Mean scores for *footing*

Clara:					
weak	fairly strong		very strong		ambiguous valence
	pos. valence	neg. valence	pos. valence	neg. valence	
<i>cGans 1</i> (+1.5)	<i>cBlum1</i> (+1.9*)	<i>cMango 1</i> (-1.9*)	<i>cSchoko 3</i> (+2.3)	<i>cKuli 3</i> (-2.2)	<i>cZwieb 2</i> (+/-2.2)
<i>cKuli 1</i> (+1.3)	<i>cMango 3</i> (+1.5)	<i>cMango 2</i> (-1.7*)			<i>cVani 2</i> (-/+2.1)
<i>cMoon 1</i> (+1.4)	<i>cGans 4</i> (+1.6*)	<i>cBlut 4</i> (-1.7*)			<i>cBlut 3</i> (+/-2.0)
<i>cZwieb 3</i> (+1.1)	<i>cBlum 3</i> (+1.9*)				<i>cSchoko 2</i> (+/-1.1)
<i>cBlum 4</i> (+1.0)	<i>cMoon 3</i> (+2.0)				<i>cKuli 2</i> (+/-1.5*)

Response distributions: *cSchoko 2*: 9: 7; *cBlut 3*:9:7; *cVani 2*:9:7; *cKuli 2*:8:8; *cZwieb 2*:8:8

Sybille:					
weak	fairly strong		very strong		ambiguous valence
	pos. valence	neg. valence	pos. valence	neg. valence	
<i>sKuli 1</i> (+1.2)	<i>sMoon 1</i> (+1.9*)	<i>sKiwi 1</i> (-2.0)	<i>sKiwi 3</i> (+2.3*)	<i>sMango 3</i> (-2.3*)	<i>sBlum 1</i> (+/-1.2*)
	<i>sMango 4</i> (+1.9*)	<i>sKiwi 2</i> (-1.7*)	<i>sBlum 3</i> (+2.0)	<i>sSchoko 1</i> (-2.1)	<i>sGans 1</i> (+/-1.7*)
			<i>sKiwi 4</i> (+2.0)		<i>sBlum 2</i> (+/-2.0)
			<i>sKuli 4</i> (+2.2)		<i>sMoon 2</i> (+/-2.1)
					<i>sSchoko 3</i> (+/-2.1)

Response distributions: *sBlum1*: mixed;*sSchoko 4*:mixed;*sGans 1*:8:7; *sMoon 2*:8:7;*sBlum2*:8:8;*sSchoko 3*:8:8

Wu:					
weak	fairly strong		very strong		ambiguous valence
	pos. valence	neg. valence	pos. valence	neg. valence	
<i>wBing 3</i> (+1.4*)	<i>wLa 4</i> (+1.9)	<i>wLizhi 2</i> (-1.8*)	<i>wLizhi 1</i> (+2.4*)	<i>wBing 2</i> (-2.1)	<i>wMan 3</i> (+/-2.1)
<i>wLizhi 4</i> (+1.4*)	<i>wNing 4</i> (+1.8)	<i>wMan 2</i> (-1.9*)	<i>wDan 12</i> (+2.1*)		<i>wLu 4</i> (+/-2.0)
	<i>wLizhi 3</i> (+1.8*)	<i>wLa 3</i> (-1.9*)	<i>wDan 4</i> (+2.2)		<i>wNing 2</i> (+/-1.3)
	<i>wMan 4</i> (+1.7)		<i>wBing 1</i> (+2.1)		<i>wDan 11</i> (+/-1.8)

Response distributions: *wNing 2*: mixed; *wDan 11*: mixed; *wLu 4*: mixed; *wMan 3*:7:7

You:					
weak	fairly strong		very strong		ambiguous valence
	positive	negative	positive	negative	
<i>yLong 3</i> (+1.3)	<i>yLu 3</i> (+1.8*)	<i>yLing 2</i> (-1.9*)	<i>yLizhi 1</i> (+2.2*)	<i>yLong 2</i> (-2.3)	<i>yMing 1</i> (+/-2.0)
<i>yMing 3</i> (+1.5*)	<i>yMan 3</i> (+1.9*)		<i>yLong 4</i> (+2.2*)	<i>yLu 2</i> (-2.1)	
	<i>yBing 4</i> (+1.9*)		<i>yLing 1</i> (+2.2*)	<i>yMan 2</i> (-2.3)	
	<i>yLizhi 4</i> (+1.9*)		<i>yMan 1</i> (+2.4*)	<i>yDang 1</i> (-2.1)	
				<i>yLa 2</i> (-2.5*)	
				<i>yLizhi 2</i> (-2.4*)	
				<i>yLong 1</i> (-2.1)	

Response distributions: *yMing 1*:8:7

Table 4 The results of the judgement tests - The mean scores for valence/strength of position

7.2.1.4 The classification of the speech samples

With the exception of the utterances with ambiguous *valence*, which were placed in a separate category, all the 74 speech samples of the German and Chinese corpus were classified in terms of type of interaction strategy as a function of their *valence* and their mean scores for *footing*/strength of position. In the determination of speaker *footing*, the scores for *footing* and strength of position were briefly taken together and their mean scores were calculated. Speech samples with positive *valence* and a very low mean score for *footing*/ strength of position of ≤ 2.3 were classified as *deferent*, while utterances with mean scores for *footing*/strength of position of >2.3 were assigned to either the *arrogant* or the *support* category, depending on their *valence*. After this classification, the scores for *footing* and strength of position were once more kept separate as this was thought to be better for the analysis of the feature profiles (Chapter 9). The final results of this classification are shown in Table 5 below, ambiguity with respect to *valence* being marked with ‘+/-’²². As Table 5 shows, among the German speech samples those with positive *valence* and scores for *footing*/strength of position between [1.8; +1.0] or [1.7; +1.1] (in *cBlum 4* and *cZwieb 1*) and [2.9; +1.6] (in *cGans 4*) were assigned to the *deference* category²³. Among Wu’s speech samples, (positive) overall scores between [1.8; +1.8] (in *wNing 4*) and [3.0; +1.4] (in *wBing 3*) were seen as signaling *deference*, while among You’s utterances, speech samples with (positive) overall scores between [2.4; +1.3] (in *yLong 3*) and [3.1; +1.5] (in *yMing 3*) were classified as *deferent*. Interestingly, as Table 5 equally shows, none of the 74 speech samples were unambiguously assigned to the *false deference* category. However, given their very weak scores for *footing*/strength of position, two utterances, *cSchoko 2* and *cKuli 2* could, given a negative prefix, potentially signal *false deference*.

7.2.2 The influence of the discourse context

In the second half of each judgement test, the speech samples were played to the judges together with their preceding discourse context. In this subsection we will briefly take a look at how judgements undertaken with and without the discourse context may be said to differ, whether differences are systematic and what they may mean. Given the fact that the introduction of the utterance context is widely regarded as a means of disambiguating utterances, it will also be of interest to see if the results of judgements made with the discourse context appear more consistent, i.e. hold lower scores for standard deviation, than judgements made without. For this analysis, a second set of mean scores for *footing*, strength of position and *valence* was calculated for all 74 speech samples, based this time on the responses given in the second half of the tests, with the discourse context²⁴. In the following discussion we will first take a look at the differences in judgements for *footing*, *valence* and strength of position. After this, we shall come to the influence of the discourse context on standard deviation²⁵.

7.2.2.1 The influence of discourse context on judgements of footing

At first sight, the results for *footing* with and without context, shown in rows ‘x~’, columns 1 and 2 of Tables A - D (Appendix II, Section 2) seem to indicate that judgements of *footing* made for one and the same speech sample with and without the knowledge of the discourse

²² Potential ambiguity with respect to *footing* or strength of position is not included in Table 5.

²³ Their mean scores for *footing*/strength of position all lie between 1.4 and 2.25. As Table 03 shows, while Clara’s *deference* group contains 6 speech samples, only one of Sybille’s utterances, *sKuli 1* (1.8), ‘qualified for’ the *deference* category. Though *sGans 1* also has a relatively low mean score for *footing*/strength of position (2.3), it had to be assigned to the ambiguous group due to its ambiguous *valence* scores. This also applies to *sBlum 1* (2.15), *cKuli 2* (2.15) and *cSchoko 2* (1.7).

²⁴ These are to be found in Tables A - D of Appendix II, Section 2. For judgements made with the discourse context, mean scores for *footing* (also in row ‘x~’) are shown in column 2, for strength of position in column 3b and for *valence* in column 4b.

²⁵ As the judgements made without and with context were separated by a time-span of about 2 hours, this might lead to the conclusion that what we are dealing with here are two entirely separate judgements made under different circumstances, whose results may therefore not be compared. However, since, as was discussed in Chapter 6, many judges were able to recall the judgements they had made in the first half of the test and compare them with those made in the second half, it does seem possible to speak of one test and consequently the influence of the context on judgements.

context differ very strongly, indeed almost erratically. However, a closer look at the results for three groups of speech samples, those with the weakest mean scores for *footing* (< 3.0), those

Clara	(20 utterances)		
Deference	Support	Arrogance	Ambiguous valence
<i>cBlum 4</i> (1.8; +1.0)	<i>cGans 4</i> (2.9; +1.6)	<i>cBlut 4</i> (3.5; -1.7)	<i>cKuli 2</i> (2.8; +/-1.5)
<i>cZwieb 3</i> (1.7; +1.1)	<i>cMango 3</i> (3.1; +1.5)	<i>cMango 2</i> (3.1; -1.7)	<i>cZwieb 2</i> (4.0; +/-2.2)
<i>cGans 1</i> (2.6; +1.5)	<i>cBlum 3</i> (3.3; +1.9)	<i>cMango 1</i> (4.1; -1.9)	<i>cVani 2</i> (3.9; +/-2.1)
<i>cKuli 1</i> (2.3; +1.3)	<i>cMoon 3</i> (3.2; +2.0)	<i>cKuli 3</i> (3.8; -2.2)	<i>cSchoko 2</i> (2.3; +/-1.1)
<i>cMoon 1</i> (2.5*; +1.4)	<i>cSchoko 3</i> (3.7; +2.3)		<i>cBlut 3</i> (3.5; +/-2.0)
	<i>cBlum 1</i> (3.8; +1.9)		
Sybille	(17 utterances)		
Deference	Support	Arrogance	Ambiguous valence
<i>sKuli 1</i> (2.4; +1.2)	<i>sMoon 1</i> (3.4; +1.9)	<i>sKiwi 1</i> (3.5*; -2.0)	<i>sSchoko 3</i> (3.4; +/-2.1)
	<i>sMango 4</i> (2.8; +1.9)	<i>sKiwi 2</i> (3.6*; -1.7)	<i>sSchoko 4</i> (3.6; +/-2.3)
	<i>sKiwi 4</i> (3.1; +2.0)	<i>sMango 3</i> (4.4; -2.3)	<i>sBlum 2</i> (3.3*; +/-2.0)
	<i>sBlum 3</i> (3.5; +2.0)	<i>sSchoko 1</i> (3.4*; -2.1)	<i>sBlum 1</i> (3.1*; +/-1.2)
	<i>sKiwi 3</i> (3.6; +2.3)		<i>sGans 1</i> (2.9*; +/-1.7)
	<i>sKuli 4</i> (3.8; +2.2)		<i>sMoon 2</i> (3.6; +/-2.1)
Wu	(18 utterances)		
Deference	Support	Arrogance	Ambiguous valence
<i>wLizhi 4</i> (2.3; +1.4)	<i>wBing 1</i> (3.0; +2.1)	<i>wLa 3</i> (3.3; -1.9)	<i>wDan 11</i> (3.2*; +/-2.4)
<i>wNing 4</i> (1.8; +1.8*)	<i>wDan 4</i> (3.3; +2.2)	<i>wMan 2</i> (3.8; -1.9)	<i>wMan 3</i> (2.8; +/-2.1)
<i>wLa 4</i> (2.7*; +1.9)	<i>wLizhi 1</i> (3.0; +2.4)	<i>wLizhi 2</i> (4.8; -1.8)	<i>wNing 2</i> (2.4; +/-1.3)
<i>wBing 3</i> (3.0*; +1.4)	<i>wDan12</i> (3.4*; +2.1)	<i>wBing 2</i> (4.5; -2.1)	<i>wLu 4</i> (2.6; +/-2.0)
<i>wLizhi 3</i> (2.6; +1.8)	<i>wMan 4</i> (3.7; +1.7)		
You	(19 utterances)		
Deference	Support	Arrogance	Ambiguous valence
<i>yLong 3</i> (2.4; +1.3)	<i>yBing 4</i> (2.9; +1.9)	<i>yLing 2</i> (3.6; -1.9)	<i>yMing1</i> (3.3; +/-2.0)
<i>yLu 3</i> (2.7; +1.8)	<i>yMan 1</i> (3.1*; +2.4)	<i>yDang 1</i> (3.3; -2.1)	
<i>yLizhi 4</i> (2.7; +1.9)	<i>yLing 1</i> (3.2; +2.2)	<i>yMan 2</i> (4.1; -2.3)	
<i>yMing 3</i> (3.1; +1.5)	<i>yMan 3</i> (3.1; +1.9)	<i>yLu 2</i> (3.6; -2.1)	
	<i>yLizhi 1</i> (3.3; +2.2)	<i>yLong 2</i> (4.3; -2.3)	
	<i>yLong 4</i> (3.3; +2.2)	<i>yLong 1</i> (3.6; -2.1)	
		<i>yLizhi 2</i> (4.6; -2.4)	
		<i>yLa 2</i> (4.4; -2.5)	

Table 5 The results of the judgement tests: The classification according to type of interaction strategy

with relatively strong scores (between 3.0 and 3.5) and those with the strongest scores for *footing* (≥ 3.6), reveals a few interesting patterns, to be discussed separately for each speaker.

Of Clara's 8 speech samples with very weak scores for *footing*, 7 become stronger with context, only one, *cGans 4*, becoming weaker, from [+2.9] to [+2.7]; Of her 8 utterances with relatively strong scores for *footing*, 6 become stronger, one, *cBlut 3*, remains unchanged, at [+3.5], and one, *cBlut 4*, becomes weaker, from [+3.4] to [+2.8]. Finally, of her six speech samples with the strongest scores for *footing*, half become stronger and half become weaker with context. Thus, it would seem that, broadly speaking, there was a general tendency to judge all of Clara's isolated utterances as stronger with context, less strong in isolation.

Of Sybille's 3 very weak utterances, 2 become stronger with context, while one, *sGans 1*, became weaker, changing from [+2.9] to [+2.8]. Of her 8 utterances with relatively strong scores for *footing*, 4 became stronger with context, three became weaker and one, *sKiwi 4*, remained the same, at [+3.1]. Finally, of her 6 speech samples with very high scores, 5 became weaker and one, *sKiwi 2*, stayed the same, at [+3.6]. Hence, though these patterns are less clear than with Clara's utterances, it seems possible to say that generally speaking Sybille's utterances, especially her strong ones, were judged stronger in isolation than with context.

Turning now to Wu, all of her 7 very weak utterances became stronger with context. Of her 6 utterances with relatively high scores for *footing*, half became stronger with context and half became weaker. Finally, of her four utterances with very strong scores, three became weaker, while one, *wMan 2*, became stronger, changing from [+3.8] to [+4.0]. In conclusion, while all of Wu's weak utterances were judged weaker without context, her very strong utterances were judged weaker with context, stronger in isolation.

Finally, all of You's 4 very weak utterances were judged stronger with context. Of her eight speech samples with relatively strong scores for *footing*, four became weaker with context, two became stronger, and two, *yMan 3* [+3.1] and *yLizhi 1* [+3.3], stayed the same. Finally, of her 7 utterances with very strong scores, 5 were judged less strong with context, one, *yLong 2* [+4.3], remained unchanged and one, *yLu 2*, became stronger, changing from [+3.6] to [+3.8]. Thus, similar to what was found with Wu, all of You's weak utterances were judged weaker in isolation, while most of her strong utterances were judged stronger without context.

To summarize, though the results often vary, especially for utterances originally²⁶ holding mid scores for *footing*, it does seem possible to say that for all four speakers the speech samples with originally weak scores for *footing* were judged stronger after the introduction of the discourse context. On the other hand, with the exception of Clara, the majority of the utterances with originally very strong scores for *footing* were judged weaker with context. Interestingly, Clara's originally strong utterances were judged even stronger with context.

7.2.2.2 The influence of discourse context on judgements of valence and strength of position

While the effect of the discourse context on judgements of these two categories appears similarly diverse, a close scrutiny of individual results shows a number of broad tendencies of the kind discussed above for *footing*. In the following the differences in judgement will be examined for the following 4 groups of speech samples: (1) speech samples originally holding weak mean scores for strength of position (< 1.8) and positive *valence*, (2) speech samples originally holding strong mean scores for strength of position (≥ 1.8) and positive *valence*, (3) utterances with initially strong scores for strength of position, but negative *valence* and (4) speech samples originally judged to be ambiguous for *valence*²⁷. As above, these groups will be examined separately for each speaker.

²⁶ What is referred to here as the 'original' or 'initial' judgement, or score, of a speech sample, or utterance, relates to the judgement it received in the first half of the judgement test, i.e. without the discourse context.

²⁷ The mean scores for valence made with and without the context are shown in row 'x~' of columns 3b and 4b, in the case of strength of position, the two kinds of results are to be found in row 'x~' of columns 3a and 4a (Appendix II, Section 2).

Similar to what was observed above for *footing*, all of Clara's 5 weak and positive utterances became stronger with context, all but one retaining their positive sign. Only *cGans 1* became negative with context. Of Clara's strong and positive utterances, 3 became stronger with context, 2 remained the same, and one, *sSchoko 3*, became weaker, changing from [+2.3] to [+2.2]. Again, *valence* remained the same in almost all cases, changing from positive to negative in only one case, *cMango 3*. Of Clara's 4 strong and negative utterances, three became stronger and one, *cBlut 4*, became weaker, changing from [-1.7] to [-1.6]. The negative sign remained in two utterances, but changed to a positive sign in one case and to ambiguous *valence* in one further case. Finally, of Clara's 5 speech samples with ambiguous *valence*, four became stronger with context, and *valence* was now either positive or negative in all cases. Therefore, broadly speaking, with respect to strength of position it appears that the majority of Clara's weak and strong utterances were judged weaker in isolation than with the discourse context. While positive *valence* generally remained positive after the introduction of the context, most cases of negative or ambiguous *valence* changed to positive *valence* in the second half of the judgement test.

Coming now to Sybille, her one weak and positive utterance became stronger with context, *valence* remaining unchanged. Of her 6 speech samples with strong scores for strength of position and positive *valence*, three became weaker with context, 2 remained the same, and one, *sMango 4*, became stronger, changing from [+1.9] to [+2.0]. The positive sign remained unchanged in all 6 cases. Of her 4 strong utterances with negative *valence* two became weaker, one, *sMango 3* [-2.3] remained the same and one, *sKiwi 2*, became stronger, its scores changing from [-1.7] to [+2.1]. *Valence* remained negative in one utterance, changed to positive undertone in two cases, and to ambiguous *valence* in one utterance, *sKiwi 1*. Finally, of Sybille's 6 utterances with ambiguous *valence*, 3 became weaker, 2 became stronger and one, *sSchoko 3*, remained the same. Ambiguous *valence* changed to positive in almost all cases, to negative *valence* in *sBlum 2*. To conclude, while the patterns here are less clear than with Clara, these findings do seem to indicate that a large number of Sybille's originally strong utterances were judged less strong after the introduction of the discourse context. The majority of her negative signs changed to positive *valence*, while in almost all cases ambiguous *valence* changed to positive *valence*.

Turning to Wu, two of her originally weak utterances with positive *valence* became stronger and remained positive with context. Of her 8 speech samples with strong scores for strength of position and positive *valence*, 3 became stronger, 3 became weaker and two, *wNing 4* and *wLa 4*, remained unchanged, at [+1.8] and [+1.9], respectively. Positive *valence* changed to negative *valence* in only one case, *wDan 4*. Of Wu's 4 strong utterances with negative *valence*, 2 became stronger with context, one, *wMan 2* [-1.9], remained the same, while one, *wBing 2*, became weaker, changing from [-2.1] to [-1.9]. Only in *wLa 3* did negative *valence* change to positive *valence*. Of her 4 utterances with ambiguous *valence*, three were judged stronger with context, one, *wMan 3*, becoming weaker, changing from [+/-2.1] to [+/-2.0]. Ambiguous *valence* changed to positive in all cases. To conclude, while no clear patterns of change are to be found with respect to strength of position, it does seem possible to say that changes from originally positive to negative *valence* were rare and that all cases of initially ambiguous *valence* changed to positive *valence* after the introduction of the discourse context.

Finally, both of You's initially weak utterances with positive *valence* became stronger and remained positive with context. Of her 8 speech samples originally given strong scores for strength of position and positive *valence*, all became stronger with context. Positive *valence* remained positive in all cases. Of You's eight originally strong utterances with negative *valence* 4 became stronger with context, two remained the same and two utterances, *yLong 2* and *yLa 2*, became weaker, changing from [-2.3] to [-2.2] and from [-2.5] to [-2.1], respectively. Negative signs changed to positive signs in only two cases, *yDang 1* and *yLong 1*.

Finally, You's utterance with ambiguous *valence*, *yMing 1*, became weaker with context and also changed to positive *valence*. Thus, in conclusion, the majority of You's originally strong utterances became weaker with context. While her one utterance with ambiguous scores of *valence* changed to positive *valence* in the second half of the test, only very few negative signs changed to positive signs after the introduction of the discourse context.

In conclusion, the above discussion revealed the following findings: First, it may be said that the majority of the German and Chinese speech samples were indeed judged differently before and after the introduction of their discourse context. The fact that so many of the speakers' utterances were interpreted differently under different circumstances shows how ambiguous the signaling of attitude can be. Second, with respect to specific patterns relating to these differences, the results of the above analysis vary quite considerably from one speaker to the next. However, a few general points may be pointed out. Thus, for example, most of the speech samples in the corpus which originally received weak scores for *footing* received stronger scores for this category after the introduction of the context. As a second general observation, while the majority of Sybille's, Wu's and You's utterances with very strong scores for *footing* became weaker in the second half of the test, exactly the opposite happened in the case of Clara, whose strong speech samples all became even stronger with context. Third, regarding judgement differences with respect to strength of position and *valence*, a large number of speech samples with weak scores for strength of position and positive *valence* remained positive but were judged stronger after the introduction of the discourse context. While the patterns with respect to the strong utterances were less clear-cut, it does appear that Sybille's and You's strong utterances were generally judged stronger in isolation, while most of Clara's strong utterances were judged weaker in isolation. Finally, with regard to *valence*, while changes between positive and negative signs varied, it seems possible to say that in the majority of cases originally ambiguous *valence* changed to either positive or negative *valence* once the context was known. Among Clara's speech samples, positive *valence* generally remained positive, and most of Clara's and Sybille's utterances with negative or ambiguous *valence* took on positive signs after the introduction of the discourse context.

7.2.2.3 The influence of the discourse context on standard deviation

We now come to the examination of the effect of the discourse context on standard deviation, i.e. a comparison of standard deviation values with and without context. These are found for each speech sample in row 'sd' of columns 1 and 2 (for *footing*), 3a and 4a (for strength of position) and 3b and 4b (for *valence*) of Tables A - D, in Appendix II, Section 2. The results of this comparison are summarized in Table 6, showing how the introduction of the discourse context affected the standard deviation of judgements of *footing*, strength of position and *valence* in the speech samples of the four speakers.

As Table 6 shows, while the influence of the discourse context on the standard deviation of *valence* judgements clearly varies strongly, the introduction of the context does seem to have lowered standard deviation with respect to *footing* in the majority of Clara's, Sybille's and Wu's utterances. This indicates that responses were more in agreement here, pointing to an overall lesser degree of ambiguity on the part of the speech samples. Interestingly, in the case of You, the trend goes in the opposite direction: Here, the introduction of the context incurred an increase of standard deviation, suggesting that the introduction of the discourse context rendered her utterances no easier - indeed, as it seems, even more difficult - to judge. Finally, for all four speakers the clarification of the discourse context incurred an overall decrease of standard variation with respect to strength of position²⁸.

²⁸ The latter trend is less marked in the case of Clara.

Clara	footing	strength of position	valence
s.d. value increases	8	7	6
s.d. value decreases	10	8	7
s.d. value unchanged	2	5	7
Sybille	footing	strength of position	valence
s.d. value increases	4	3	5
s.d. value decreases	8	11	5
s.d. unchanged	5	3	7
Wu	footing	strength of position	valence
s.d. value increases	6	4	2
s.d. value decreases	9	12	7
s.d. unchanged	3	2	9
You	footing	strength of position	valence
s.d. value increases	10	3	7
s.d. value decreases	5	11	7
s.d. unchanged	4	5	5

Table 6 The influence of the discourse context on standard deviation

7.2.2.4 Summary: the influence of the discourse context

The above discussion has yielded the following results: For the vast majority of the German and Chinese speech samples the introduction of the discourse context incurred changes in the judges' interpretations of speaker attitude, in terms of the interaction strategy thought to be signaled in each case. The finding that so many of the speakers' utterances were interpreted differently under these changing circumstances, not only with respect to *footing* and strength of position but also with respect to *valence*, their undertone, may be taken as an indication for the high degree of ambiguity manifested in the communication of both (*footing*/strength of position and *valence*) aspects of interpersonal attitude. With regard to the precise effect of the context on individual judgements, this was found to vary quite considerably, both between speakers and among the speech samples of one speaker. A few general statements are, however, possible: Generally speaking, weak scores for both *footing* and strength of position became stronger after the introduction of the discourse context, while originally strong speech samples became weaker in the second half of the test. As was observed above, the exception to this latter pattern is Clara, whose weak *and* strong speech samples all received stronger scores after the introduction of the discourse context. The fact that the large majority of Clara's utterances were judged weaker in isolation than with context could conceivably be interpreted as indicating that, in contrast to Sybille's speech samples, for example, where this effect was not observed, Clara's utterances contained some specific feature or features, possibly of a suprasegmental nature, which led the judges to give all of them relatively weak responses, as compared to Sybille's²⁹. With respect to *valence*, it seems possible to say that

²⁹ This issue was not closely examined in the present study. However, it may be said that Clara's voice, as compared to Sybille's, is relatively breathy, and that she has a strong tendency to use *palatalization*.

while changes between positive and negative signs were rare, most of the utterances which had been judged as ambiguous in isolation were interpreted as signaling either positive or negative *valence* once the context was known. This may be interpreted in terms of a disambiguation of originally ambiguous speech samples.

Regarding the effect of the discourse context on standard deviation, it was found that standard deviation clearly decreased for three speakers, Clara, Sybille and Wu, with respect to judgements of *footing* and strength of position. This suggests that the introduction of the discourse context had the effect of disambiguating the speech samples of these speakers with respect to these two components of the metamessage. Interestingly, no such effect could be found with respect to You, whose speech samples seemed no easier - indeed possibly even more difficult - to judge with the context. The reason for this could not be immediately clarified. Finally, the effect of introducing the discourse context on judgements of *valence* appears to have varied very strongly in all cases, indicating that generally speaking, the introduction of the discourse context did not make judgements of the undertone any easier and may in fact have rendered the judges' task even more difficult, by changing or increasing the information load which they had to process. However, it was found that speech samples whose valence had been impossible to judge without context were indeed disambiguated by the introduction of their context.

7.2.3 Summary and conclusions

The results of the sociolinguistic judgement tests and the findings of the quantitative evaluation of the judges' responses may be summarized as follows: First, of the 103 German and Chinese speech samples offered in the judgement tests, 74 (71.8%) were analysed in a satisfactory manner with respect to interpersonal attitude, i.e. interaction strategy. A first evaluation of the judges' responses revealed 76 cases of slight to strong inconsistencies in the judgements relating to *footing*, strength of position or *valence*, suggesting ambiguities with respect to these components. These results show that the assessment of speaker attitude on the basis of speech alone is far from easy and involves the same difficulties and ambiguities as have been reported for English. On the other hand, a closer examination of the results, this time based on response distributions, showed that really strong discrepancies in the responses were present in only 28 cases, involving a total of 25 speech samples³⁰.

As the second result, the results of the quantitative analysis showed that all but 16 of the 74 speech samples could be assigned to one - in cases of ambiguity maximally two - of the interaction strategies *support*, *deference* and *arrogance*, a fact that indicates that these interaction strategies are indeed known in Chinese and German and are part of interpersonal communication in these languages. On the other hand, only two German utterances were judged in terms of *ambiguous*, and potentially *false*, *deference*, none of the speech samples unambiguously assigned to this fourth category. This may be regarded as evidence for the strong conventionalization of *deference* in Chinese society, due to which potentially ambiguous deferent Chinese utterances are nevertheless generally classified as deferent. On the other hand, it could also have been that the judges in the tests shrank from classifying any of the speakers' utterances in terms of this extremely negative speech behavior.

As a third result, the examination of the responses given in the second half of the judgement test revealed that the majority of the German and Chinese speech samples were judged differently in context than in isolation, statements to this effect made by the judges themselves after the completion of the tests. This may be regarded as evidence of the ambiguity of human speech with respect to interpersonal attitude. Moreover, while the majority of Sybille's, Wu's

³⁰ In the determination of the phonetic exponency of the strategies *support* and *arrogance*, *deference* and *false deference* in Chapter 9, discussions will center on discrepancies (i.e. ambiguity) with respect to *valence*. As the role of *footing* plays no role in the differentiation between these strategies, the ambiguities found for *footing* and strength of position will not receive further consideration here.

and You's speech samples were given weaker scores for *footing*/strength of position in the second half of the test, the opposite trend was observed for Clara's utterances, which received stronger scores in context. This suggests that Clara's speech, or her voice, contained features which led the judges to interpret her speech behavior as gentle and weak. Finally, while the introduction of the context clearly had no effect on *valence* judgements, it was observed that in the majority of the speech samples the assessment of footing and/or strength of position seems to have become easier with the disclosure of the discourse context.

8. The phonetic microanalysis

The objective of the phonetic microanalysis was to determine the suprasegmental features of the 74 German and Chinese speech samples classified according to interaction strategy in the sociolinguistic judgement tests. The results of the analysis, in terms of the phonatory and supralaryngeal *settings*, the overall pitch, loudness, tempo and fluency, and intonation of these utterances, were compiled in the form of the *suprasegmental profiles*¹, shown in Table 10. They were to serve in the determination of the phonetic exponency of the interaction strategies, discussed in Chapter 9.

The phonetic microanalysis consisted of three stages. As the first step, the 74 speech samples were analysed auditorily in analytic listening tests with linguistically non-naïve participants. Following this, the results of these analyses were verified and complemented with the help of instrumental speech signal analyses. And as the third and final stage, the results of the auditory and instrumental analyses were brought together to obtain the final results of the microanalysis, the *suprasegmental profiles*.

8.1 The analytic listening tests

The discussion of the analytic listening tests begins with the general procedure of the tests, after which the participants, the speech material and the questionnaires are briefly presented.

8.1.1 General procedure

The general procedure of the analytic listening tests, hereafter referred to as the listening tests, was largely identical to that of the sociolinguistic judgement tests. As before, small groups of Chinese and German judges were asked to analyse the speakers' utterances with the help of a questionnaire. This time, however, the judges were less numerous as they had to be linguistically non-naïve. The speech material was also more limited, consisting only of the 74 speech samples which had been classified in the sociolinguistic judgement tests. As a third difference, the participants in the listening tests did not all share the same task. Instead, the tests were divided into three parts, each pertaining to one set of suprasegmental features and assigned to one group of judges. While the German listening tests were held at the Phonetics Institute at the University in Saarbrücken, the Chinese tests were held at the home of one of the judges. Each listening test took between 45 and 60 minutes.

8.1.2 The participants

In contrast to the sociolinguistic tests, the analytic listening tests necessitated a certain degree of linguistic knowledge or at least awareness on the part of the judges. Thanks to the participation of the students and lecturers of the Phonetics Institute in Saarbrücken, judges for the German tests were easy to find. The auditory voice quality analysis of all the German speech samples was accomplished by three lecturers of the Institute, and the tests for overall pitch, loudness, tempo and fluency were performed by one lecturer and two advanced students of the Institute specializing in this field. Finally, the analysis of German intonation was under-taken by the experimenter and one lecturer of the Institute². Thus, the German listening tests involved a total of eight judges, none of whom, apart from the experimenter herself, was acquainted with the speakers Clara and Sybille.

Finding participants for the Chinese listening tests was more difficult. As the auditory analysis of voice quality does not require any specific knowledge of the grammar and phonology of the language concerned, the voice quality analysis of the Chinese speech samples was performed

¹ The *suprasegmental profiles* are shown in Tables 10 at the end of this chapter.

² Following separate analyses, individual transcriptions were compared and discussed, after which a final analysis of each German utterance was agreed upon.

by the three lecturers who accomplished the German voice quality tests. For the auditory analysis of the overall features, however, native speakers of Chinese were required and great effort was taken to find six linguistically non-naïve Chinese native speakers. After a careful selection among the Chinese students she was acquainted with, the experimenter chose three from the Peoples' Republic and three from Taiwan. While only one of these had received formal training in linguistics, a second was a lecturer of Chinese at the College of Adult Education in Saarbrücken. All 6 judges displayed a high degree of linguistic awareness which was deemed adequate for the accomplishment of the tests³. The judges from Taiwan were all well acquainted with speaker You, but only one of the Mainland Chinese judge knew Wu personally. The analysis of Chinese intonation was undertaken by the experimenter, based on the model of Mandarin Chinese intonation presented in Chapter 3 and following the methodology to be outlined below. The auditory analyses of the Chinese speech samples were thus accomplished by a total of ten judges: the experimenter, the 3 lecturers of the Phonetics Institute and the 6 native Chinese participants.

8.1.3 The speech material and the questionnaires

The speech material of the listening tests consisted of the 74 German and Chinese speech samples which had been classified according to interaction strategy. Similar to the procedure followed in the sociolinguistic judgement tests, these were played to the judges in randomized order. This time, however, the discourse context was not included and the speech samples were offered to the judges in isolated form only. As in the sociolinguistic tests, the speech samples of the respective speaker were played to the judges in a block before the beginning of each test, so that they could familiarize themselves with her voice and gain a first impression of the phenomena they would be asked to determine. During the tests each speech sample was offered a minimum of 6 and a maximum of 9 times.

As each test focused on a different set of suprasegmental features, each group of judges received a different questionnaire. In the questionnaire used in the analysis of voice quality, the judges were asked to indicate the voice quality *setting(s)* of each utterance, or subsections thereof, choosing from the following supralaryngeal and phonatory *settings*: *palatalization (fronted speech)*, *backed speech*, *modal voice*, *tenseness*, *laxness*, *breathiness*, *whisper* and *creak*⁴. For the analysis of Chinese voice quality a representation of the Chinese speech samples using the *pin-yin* transliteration system was included in the questionnaire⁵. The analysis of the overall features was divided into two parts, each consisting of two assignments. In the first part, the judges were asked to indicate overall tempo, choosing between (*very*) *fast*, *normal* or *mid* and (*very*) *slow*, and to give the locations and relative lengths of marked pauses and lengthenings, the term 'marked pause/lengthening' being used to denote non-phonological pauses and lengthenings and phonological pauses/lengthenings of overly long duration⁶. In the second part of the test the judges were asked to indicate overall loudness and pitch, choosing between (*very*) *high/ loud*, *normal/mid* and (*very*) *low/soft*. The questionnaire used in these tests contained orthographic (in German) and transliterated (in Chinese) representations of the speech samples, below which responses were to be indicated, either in words or with the help of schematic drawings. For maximum clarity and as an

³ Above all, these six students had acquired a high proficiency in the German language which, given the difficulties of German phonology, morphology and syntax for Chinese native speakers, was seen as an indication of their above-average linguistic awareness. (See Gao 1983). Regrettably, it was not possible to enlist the assistance of the Chinese lecturer of the university. The former lecturer who had participated in the recordings had since left and could not be reached either.

⁴ The questions of this questionnaire are shown in Section 1.2 of Appendix III.

⁵ See Section 2.2 of Appendix I. The *pin-yin* transliteration system is presented in Section 1 of Appendix I.

⁶ The term 'marked' was explained to the judges as relating to pauses and lengthenings whose location in utterances is 'unexpected' or 'unusual', giving rise to a 'strange', or 'unusual' auditory impression, or feeling. The questionnaire used in these tests is shown in Section 1.2. of Appendix I.

illustration of their responses, the participants were asked to divide each utterance into subsections, as appropriate. While the questionnaires for all the listening tests were written in German only, in the tests with the Chinese native speakers the questions were translated and discussed in Chinese to ensure that they were fully understood. The German intonation analysis focused on the determination of the *nuclear* and *onset* syllables, the type and relative pitch-range of the nuclear pitch-patterns, the pitch-patterns in the *heads* and the configuration of the *tail* slopes. The analysis of Chinese intonation centered on the segmentation of utterances into units of *tonality*, i.e. (major and minor) intonation-groups, tone-groups and tone-units, and the determination of prominent/stressed and focused syllables.

8.1.4 The results of the listening tests

The analytic listening tests went very smoothly, only a few participants having occasional difficulties in accomplishing the tasks. In the voice quality tests, which were held in the form of discussion rounds, the judges' opinions sometimes differed with respect to the phonatory *settings* of German or Chinese utterance. Interestingly, no supralaryngeal *settings* were detected in any of the examined speech samples⁷. With regards to the tests for the overall features, the examination of the results showed that one Chinese judge had at times indicated the presence of both phonological and non-phonological pauses/lengthenings, while a second had not been able to accomplish the task for overall pitch-height. While the 'errors' of the first judge were negligible⁸, the responses of the second Chinese judge for pitch-height could not be used and this part of the test was therefore repeated with a different judge, also from China, great care being taken to familiarize him with the speaker's voice before he answered the questions⁹. The results of the analytic listening tests are incorporated in Table A of Appendix III (Section 2), for the phonatory *settings*, and B, for the overall features. The results for intonation are shown in Table C, those for the supralaryngeal *settings* in Table D¹⁰.

8.2 The instrumental speech analyses

The first, main purpose of the instrumental speech analyses was to verify the results of the listening tests. A second objective was to examine three phenomena not included in the tests: (1) Overly short or missing phonological pauses and lengthenings, (2) degree of utterance-final slow-down and (3) the presence of speech errors¹¹.

8.2.1 Procedure and methodology of the instrumental analyses

In preparation for the instrumental speech analyses, the 74 German and Chinese speech samples were digitized at a sampling rate of 16,000 Hz and copied onto an optical (MO) disc. Before digitization the amplitudes of the speech samples were adjusted to avoid over- and undermodulation. After a brief, preliminary examination of the digitized speech samples during which, among other things, their segmentation into subsections was verified, the instrumental speech analyses were carried out with a Kay Elemetrics CSL speech analysis system, in the manner to be outlined below.

⁷ It was concluded that the detection of supralaryngeal *settings* obviously necessitates a high degree of familiarity with the speaker's voice, which the judges in the listening tests did not possess.

⁸ The 'errors' of this person were negligible as his responses only occasionally differed from those of the other two judges. The term 'error' is between quotation marks to illustrate that what we dealing with here are not *mistakes*, but the judge's inability to distinguish clearly between phonological and non-phonological pauses, a process which is considerably more complex in Chinese than in non-tone languages,

due to the stronger variations in tempo and the greater number of tonality units found there.

⁹ The original participant was not asked to repeat the listening test out of consideration for his *face*. In the representation of the results the data which relate to the responses of these two different judges are marked with a '*'.
¹⁰ The symbols and notation marks used in these tables are presented in the introduction to section 2 of Appendix III.

¹¹ The auditory analysis of these features in the German speech samples was carried out by the experimenter. The presence of pauses and lengthenings and the occurrence of marked utterance-final slow-down in the Chinese speech samples were determined with the help of discussion rounds involving small groups of some of the Chinese participants in the listening tests.

8.2.1.1 The instrumental analysis of phonatory voice quality

The instrumental analyses of phonatory voice quality were carried out with speech pressure waveforms, synchronized laryngeal waveforms and spectral sections, and performed in two steps for stressed/prominent vowels¹². First, the primary mode of phonation, i.e. *modal voice*, *tenseness* or *laxness*, was determined as an indicator of overall tenseness in the articulatory system. This was done by calculating the mean *closed quotient* (*CQ*), which involved calculating the *CQ* (indicated in %) at three locations of the vowels concerned and determining their mean value, also in %¹³. The second step involved distinguishing between *lax* phonation with and without concomitant air-turbulences, this being a prerequisite for the distinction between *laxness* and *breathiness*. Here, the intensities of the 1st and 2nd harmonics of the vowels concerned were measured and their difference (in Hz) calculated, results of ≥ 12 dB taken to indicate *breathiness*. Since this method is not applicable to high close vowels, such as [i], [u], [y] and [ɨ]¹⁴, in the examination of such vowels the spectrograms were examined for air-turbulences not originating in the articulation of consonants immediately preceding them¹⁵.

8.2.1.2 The instrumental analysis of the overall features

In these analyses the following features were determined, either for each utterance as a whole or for individual subsections:

- (1) overall/mean frequency, intensity and tempo (in terms of articulation-rate)
- (2) the location and duration of phonological and non-phonological pauses/lengthenings, and the absence of phonological pauses/lengthenings
- (3) the presence of speech errors
- (4) the utterance-final slow-down rate

Overall/mean intensity, obtained from the intensity contours of utterances, was determined for each speech sample as a whole, results being indicated in dB. and in relative terms¹⁶. Overall tempo, i.e. articulation rate, was calculated with the help of spectrograms in terms of mean segment duration (in ms/segment) and mean syllable duration (in msec/syllable) per subsection¹⁷. The method of calculation of overall/mean frequency, achieved with the help of *fo*-contours, differed slightly in the German and Chinese speech samples. In the German speech samples overall frequency was indicated in terms of the frequency of the functionally

¹² In German, as in English, stressed/prominent vowels and unstressed/non-prominent vowels differ in a number of aspects: Non-prominent vowels are shorter and laxer than prominent vowels and are also marked by more centralized vowel quality. In order to ascertain as constant a vowel quality as possible, all non-prominent vowels were excluded from these calculations, a measure also applied to the Chinese material. While the differences in duration are not as marked in Chinese as in German or English, quality differences between stressed/prominent and unstressed/non-prominent vowels are also strong here.

¹³ The *CQ* was determined by dividing the duration of the closing phase by the duration of the entire of the entire glottal cycle, results being rounded up to one decimal point. Based on the results of the auditory listening tests and a careful auditory evaluation of each speech sample, a *CQ* of $\leq 31.9\%$ was taken as representing *laxness*, a *CQ* of 32-33.9% as *modal/laxness*, a *CQ* of 34%-37.9% as representing *modalness*, a *CQ* of 38%-40.9% as *modal/tenseness* and a *CQ* of $\geq 41\%$ as representing *tenseness*.

¹⁴ This method of calculation is not applicable to high close vowels due to overlaps of low first formant frequencies and first harmonics. As the first harmonic corresponds to fundamental frequency, such overlaps are particularly frequent in female speech. In the case of closing diphthongs, such as [au] and [aɨ], the calculations were carried out for the non-high vowels.

¹⁵ The effect of air-turbulences from preceding fricatives and aspirated plosives ([p^h], [t^h] and [k^h]) is quite strong in German and Mandarin. Mandarin Chinese also has strongly aspirated affricates ([tʃ^h], [tʃ^h] and [tʃ^h]), represented in *pin-yin* as /c/, /ch/ and /q/.

¹⁶ In the determination of relative measures, intensity values above 72 dB were taken as indicative of *high* mean intensity, values between 58 and 72 dB as *mid* and values below 58 dB were taken representing as *low* intensity. This evaluation system was devised on the basis of a conscientious comparative evaluation of the results of the listening tests and the auditory impression of each speech sample, the results of which were brought together with the results of the instrumental analyses.

¹⁷ Mean syllable and segment duration were calculated by dividing the length of each subsection, excluding silent periods, by the number of syllables and segments contained therein. Syllable reductions were dealt with by taking the underlying phonemes. In the determination of relative measures for overall/mean tempo, values ≤ 63 ms. were taken as *fast*, values between 64 and 97 ms. as *mid*, values between 98 and 124 ms. as *slow* and results ≥ 125 ms. as representing *very slow* tempo. This evaluation system was devised in the same manner as described above for the relative measures of overall/mean intensity.

important syllables of each intonation-group¹⁸. Overall/mean frequency in Chinese was indicated in terms of the *bottom line* of speech samples, corresponding to the lowest frequency of the 3rd tone(s) in each tone-unit¹⁹. Results were indicated in Hz. and in relative terms²⁰. Coming now to points (2) and (3), the occurrence and duration (in Hz) of non-phonological pauses/lengthenings and phonological pauses/lengthenings as well as the incidence of speech errors was determined with the help of speech pressure waveforms and synchronized wide-band spectrograms²¹. Finally, in stage (4), the degree, or rate, of utterance-final slow-down was determined for certain utterances of interest, i.e. with the marked auditory feeling towards the end, by relating the mean segment durations in their ultimate and penultimate subsections to each other²². For this, the mean segment length of the slower subsection, in most cases the ultimate one, was divided by that of the faster subsection, normally that of the penultimate section, the result taken to represent the slow-down rate²³. In cases where the mean segment duration of the ultimate subsection was found to be higher than that of the penultimate section - a phenomenon reflecting utterance-final 'speeding-up' - the results of the calculations were prefixed with a negative sign ('-').

8.2.1.3 The instrumental analysis of intonation

The instrumental analysis of German intonation

For the analysis of German intonation, the results of the listening tests relating to the location and type of nuclear pitch-patterns and the configurations of *heads* and tail-slopes, were verified, and if necessary adjusted, with the help of speech pressure waveforms and *fo*-contours. The precise frequency-range of all nuclear (accent-range) and prenuclear pitch-patterns was determined, also using *fo*-contours. The results of calculations were indicated in Hz. and in relative terms²⁴.

The instrumental analysis of Chinese intonation

The instrumental analysis of Chinese intonation, which was also based on the communicative forms presented in Chapter 3, consisted of two steps. First of all, the results of the auditory analysis were verified with the help of *fo*-contours, involving the segmentation of the utterances into units of *tonality* and the determination of prominent/stressed and focused syllables. Following this, the following three phenomena were determined using *fo*-contours:

- (1) Degree of declination
- (2) Overall (mean) tonal band-width

¹⁸ The functionally important syllables of the German utterances are *onset* and *nuclear* syllables. The pitch of syllables coinciding with the starting and ending-points of consecutive 'waves' in stepping and glissando *heads* was also determined (See Chapter 2).

¹⁹ See Chapter 3. In the absence of 3rd tones, the ending points of full, non-reduced 4th tones were taken as rough equivalents of the pitch of the *bottom line*.

²⁰ In the determination of overall/mean pitch/frequency, values ≤ 150 Hz were taken to represent *low* frequency, values between 150 and 199 Hz as *mid-low*, between 200 and 220 Hz as *mid*, between 220 and 250 Hz as *mid-high* and values above 250 Hz were taken as *high* frequency. This evaluation system was devised in the same manner as described above for the relative measures of overall/mean intensity.

²¹ Speech errors occurred a total of four times, in *sSchoko 1*, *cZwieb 3*, *yLizhi 2* and *wLizhi 3*. More will be said on this subject below.

²² As described in Chapter 2, utterance-final, or preboundary, slow-down is the generally observed tendency of speakers to reduce speech tempo over the last two or three syllables of intonation-groups. The method of calculation applied here was devised in order to minimize the influence of *phonologically* induced syllable-lengthening, as in nuclear and stressed/prominent syllables, at the end of utterances.

²³ Thus, based on these calculations, a slow-down rate of 2.0 signifies that the mean segment duration of the ultimate subsection is twice as long as that of the preceding one. This in turn means that the overall/mean tempo of the ultimate subsection is twice as slow as that of the penultimate section.

²⁴ Indications of relative height of onsets and nuclear syllables followed the system given above for the determination of overall pitch/frequency. For the relative range of nuclear and prenuclear pitch-patterns, frequency values ≤ 48 Hz were taken as *narrow*, values between 48 and 59 Hz as *mid*, values between 60 and 87 Hz as *mid-wide*, values between 88 and 179 Hz as *wide* and frequency values above 188 Hz were referred to as *very wide*. This system was devised in the same manner as described above for the relative measures of overall/mean intensity.

(3) Final pitch-/frequency-range

(4) Final tonal behavior, both of the last tones/particles and of final tone-units as a whole.

(1) *Degree of declination*

The degree of declination of a Chinese utterance was defined in Chapter 3 as the difference between the pitch of a given tone at the beginning and the end of the utterance. Degree of declination in Chinese is determined on the basis of the *bottom-line*. Therefore, calculations started by determining the *bottom-line* for the first and last tone-units of each utterance. Following this, the frequency difference between these two manifestations of the *bottom-line* was calculated and the result was divided by the number of tone-units contained in the utterance²⁵. Since the course of the *bottom-line* in a given tone-unit is reflected by the lowest frequency attained by its 3rd tone(s), degree of declination was calculated in this manner, whenever possible. In tone-units devoid of 3rd tones, calculations were based on the lowest frequency attained by full, unreduced 4th tones or the frequency of the lowest neutral(-ized) tone, as available²⁶. In utterances where the *bottom line* is interrupted by *up-* or *downshifts*, calculations were carried out for that part with the highest number of 3rd tones. No calculations were, however, possible in speech samples where *up-* and *downshifts* strongly distort the course of the *bottom line*. Finally, as the tonal relationship between the *bottom line* and the frequency of utterance-initial exclamations is as yet unclear, tone-units containing such exclamations, viz. *oh*, *ah* and *eijo*, were excluded from calculations²⁷.

(2) *Overall/mean tonal band-width*

The overall/mean, tonal band-width of the Chinese utterances was calculated by determining the tonal band-width of (ideally) every tone-unit in this utterance and calculating their mean value, in Hz. This involved determining the *top* and the *bottom lines* of (again ideally) every tone unit and calculating their difference in frequency. The *bottom-line* of each tone-unit was determined following the method described above. As the course of the *top line* is traced by the (initial) pitch/frequency of 1st and 4th tones, the determination of the *top line* was based on these two tones, as available. In tone-units where neither of these tones were present, the highest frequency attained by non-neutralized 2nd tones were taken as approximate values of the *top-line*. In a few rare cases, the frequencies of a neutral(ized) 1st or 4th tone was taken, on the condition that they represented the highest frequencies in the respective tone-unit. Finally, in those cases where none of the above methods could be applied, the frequency-range covered by prominent 4th tones was taken as approximate values for tonal band-width.

(3) *Final pitch-/frequency-range*

The frequency-range of the final tone-unit, corresponding to its tonal band-width, was calculated by establishing its *top* and *bottom lines* in the manner described above and calculating their difference in Hz. In final tone-units containing changes in tonal band-width caused by *up-* or *downshifts* of the *bottom line*, calculations were carried out for both sides of these shifts, if possible, and the results of all calculations were indicated, separated by a '/'. Final particles were excluded from these calculations.

(4) *Final tonal behavior*

²⁵ This method of calculation was devised to take into account the fact, reported in the literature, that degree of declination is determined to some degree by the length of an utterance and the number of stressed syllables contained therein (See Chapter 3).

²⁶ This auxiliary method was devised in the absence of a calculation method for determining the bottom line/frequency of a 3rd tone of a given tone-unit on the basis of other lexical tones present. Naturally, results obtained for declination - and (mean) tonal band-width - based on lexical tones other than the 3rd (for declination) and the 3rd, 1st and 4th (tonal band-width) are approximate values, marked with

a '*' and the symbols '≤' or '≥', as appropriate.

²⁷ This also applies to the utterance-initial *na* (i.e. *namma*) found, for example, in *yLa 2*.

The final tonal behavior of a Chinese utterance was defined in Chapter 3 as the pitch-pattern traced by the (initial) pitch/frequency of utterance-final tones. Two kinds of final tonal behavior are differentiated here: (1) the *caudal tonal behavior* of the very last utterance-final tone, and (2) the *focal tonal behavior* of all utterance-final tones from the last stressed/focused syllable onwards. These two phenomena were determined as follows: The final tonal behavior of last utterance-final tones, including those of sentence particles, was determined by calculating the difference between their initial frequency and that of the preceding tone, results given in Hz. and ‘tones’ (‘T’), the latter corresponding to 20Hz²⁸. In contoured utterance-final tones, i.e. 3rd, 2nd and 4th tones and final particles bearing pitch-slopes, the ranges covered by their pitch-movements were also included²⁹. The second type of tonal behavior was described in terms of the tonal behavior of all the tones of interest, determined in the manner described above and also given in Hz. and ‘T’.

8.2.1.4 The instrumental analysis of supralaryngeal voice quality

Supralaryngeal voice quality was analysed instrumentally for ten German and Chinese speech samples which, according to the auditory judgement of the experimenter, contained a certain number of words or morphemes describable in terms of a ‘gentle’, ‘soft’ or ‘petite’ feeling³⁰. Since their auditory impression involved was highly reminiscent of the manner in which *palatalization* and *labiodentalized voice* are described in Laver 1980, analyses focused on these *settings*. Using speech pressure waveforms, wideband spectrograms and spectral sections, the formant frequencies/energy maxima of segments of the affected syllables were first determined, emphasis being placed on the frequency of F2, which is generally raised in the presence of these *settings*³¹. Following this, the formant frequencies of identical segments also produced in the recordings but unaffected by *palatalization* and *labiodentalized voice* were established and compared with those obtained for the utterances under examination³².

8.3 The final results of the phonetic microanalyses

At this third stage of the phonetic microanalyses, the results of the listening tests and the instrumental analyses were brought together to obtain the final results of the microanalyses, forming the *suprasegmental profiles* of the 74 German and Chinese speech samples, shown in Tables 10. While the instrumental results for intonation and supralaryngeal voice quality could be directly adopted as final results, since no auditory tests had been held for these phenomena, the auditory and instrumental results for phonatory voice quality and the overall features, not completely identical in all cases, had to be evaluated carefully and unified. The present section deals with the kind of differences sometimes found between the auditorily and instrumentally obtained results for these features and the manner in which these differences were dealt with to obtain the final results of the phonetic microanalysis.

8.3.1 The final results for phonatory voice quality³³

Generally speaking, the results of the auditory and instrumental analyses for phonatory voice quality were found to be in good agreement, though not in all cases identical. On the whole, the instrumentally obtained results are more differentiated than the auditory results: Minor, and usually brief, fluctuations in phonatory voice quality detected in the instrumental analyses

²⁸ The ‘tone’ is employed here as a larger, more readily appreciable unit of pitch-range than Hz. It is written with a capital ‘T’ to avoid confusion with the Chinese lexical tones.

²⁹ A single initial frequency was determined in the case of utterance-final neutral and neutralized tones.

³⁰ These were cGans 1, cMoon 1, cGans 4, cMango 4, yMing 3, yLong 1, yLing 1, wNing 4, yLizhi 4 and yBing 4

³¹ As observed in Chapter, F2 is most strongly affected by these supralaryngeal *settings* and has the additional advantage of usually being most clearly visible in spectral sections.

³² In the selection of this comparative speech material, great care being taken to ensure that the phonetic environment of the examined segments was identical in all cases.

³³ In this chapter, as in Chapters 9 and 10, *modal voice* is also referred to as *modalness*, *lax voice* as *laxness* and *tense voice* as *tenseness*.

and included in the instrumental results are often not included in the auditory responses. Sometimes, the auditory and instrumental results were found to differ in absolute terms but agreed with respect to overall tendencies. Thus, for example, a few speech samples judged auditorily to contain a development from *tense* voice quality to *modal-tenseness* were found in the instrumental analyses to manifest a change from *modalness* to *modal-laxness*. While such differences may also have originated in slight differences between the evaluation systems applied in the two types of analyses³⁴, on the other hand the objective auditory evaluation of a totally unfamiliar voice quality is a relatively difficult task, so that occasional discrepancies are to be expected here³⁵. In both the above cases the instrumentally obtained results were preferred to the auditory results. As a third point, the examination of the results also suggested that the judges had had occasional difficulties distinguishing between the auditory effects of *laxness* and *breathiness*, which are very similar. This is likely to have led to differences between the instrumental and auditory results involving these two *settings*³⁶. As a fourth point, also relating to *breathiness*, its auditory effect also appears to have been occasionally confused with that of air turbulences of non-phonatory origin, produced during the production of aspirated plosives, fricatives and, especially in Chinese, affricates³⁷. In these two cases, both involving *breathiness*, the instrumentally obtained results were preferred to the auditory results. The fifth type of difference between the auditory and the instrumental results is linked to the phenomenon that relative *tenseness* is often, though not always dependent on pitch-height. Thus, in a number of cases, auditory judgements of *tenseness* indicated by the judges on the basis of pitch-rises could not be confirmed instrumentally³⁸, while in other utterances the instrumental results clearly point to a tensing of the vocal folds not detected by the judges³⁹. Conversely, a few low-pitched syllables were instrumentally found to be *laxer* than their surroundings, a feature not detected in the listening tests. Finally, on a number of stressed/prominent syllables an increase in tenseness was detected instrumentally, which, again, was not in every case reported in the auditory tests⁴⁰. Once more, the instrumental results were taken as the final results of the microanalysis.

While the large majority of the differences between the auditory and instrumental results for phonatory voice quality are explainable in terms of the six phenomena discussed above, there is, however, one speech sample where the differences do not appear to have originated in any of these phenomena. Whereas all but the very end of *cBlum 4* was auditorily judged to be *breathy*, indeed at times even *whispery*, this could not be confirmed instrumentally, particularly not for the first word, 'Blumen'; Though the spectrogram showed clear evidence of air-turbulences in the speech signal, in all cases the calculated *CQ* for [u:] was above 38%, ruling out the presence of voice qualities other than *tenseness*⁴¹. Since air-turbulences linked to the production of fricatives and affricates may also be excluded, the origin of these turbulences could not be definitely clarified. Due to the strong auditory effect of *breathiness* in this speech sample, the auditory results were adopted in this case. The final results for phonatory voice quality are given in Section 1 (line 1) of Table 10⁴².

³⁴ In fact, the evaluation of the results showed that the auditory evaluation system applied by individual judges also often differed.

³⁵ See Pützer & Barry 2003 for a discussion of the accuracy of auditory voice quality judgements. Discrepancies of this kind are found in *wLa 3*, *sKiwi 3*, *sBlum 1*, *yMan 2*, *yLong 1*, *yLa 2*, *cGans 1*, *cMoon 1*, *cMango 3*, *cKuli 3* and *cMango 2*.

³⁶ Discrepancies of this kind are found e.g. in: *sMoon 2*, *sSchoko 4*, *wMan 2* and *wLizhi 4*.

³⁷ In such cases phonatory *breathiness* as indicated by the judges could not be confirmed by the instrumental analyses. Such differences, which mostly affected only single syllables, were found in *wLizhi 3*, *wBing 2*, *wNing 2*, *wMan 2*, *wBing 3*, *wDan 12*, *yBing 4*, *yMing 3*, *yLong 3*, *yLu3*, *yLong 4*, *yLizhi 2*, *yMan 1*, *sKuli 1*, *sMango 4*, *sSchoko 1*, *sBlum 1*, *yMing 1*, *cBlum 1* and *cVani 2*.

³⁸ These were *sSchoko 1*, *cBlum 3*, *wDan 12* and, to a certain extent, *sKiwi 2*.

³⁹ These utterances were *cMango 2*, *cBlut 3*, *sBlum 1*, *sKuli 1*, *yLu3*, *yLing 1*, *yLong 1* and *yLing 2*.

⁴⁰ These were *cKuli 2*, *sBlum 1*, *sKuli 4*, *yLing 1*, *yMing 1* and *yMan 1*.

⁴¹ It is conceivable that this auditory effect was created by an approachment of the speaker to the microphone. Such sudden movements were inevitable, as the speakers were told to act naturally and may therefore - especially the very lively Clara and You - be assumed to have moved during the dialogs. On the other hand, air-turbulences caused by a slight opening of the cartilaginous glottis (i.e. between the arytenoid cartilages), which would have been quite independent of the degree of *tenseness* in the vocal folds, are also a possible origin.

⁴² The symbols and notation marks employed in Table 10 are the same as those used in Section 2 of Appendix III, showing the results of the

8.3.2 The final results for supralaryngeal voice quality

The speech samples *cGans 1*, *cMoon 1*, *cGans 4*, *cMango 4*, *yMing 3*, *yLong 1*, *yLing 1*, *wNing 4* were examined for *palatalization*, *yLizhi 4* and *yBing 4* were examined for *labiodentalized voice*. As shown in Table D of Appendix III (Section 2), strong evidence for *palatalization* was found in the German speech samples examined for this *setting*: In all these cases, the second formant frequency of the segments under examination was higher than in the comparative speech material. Coming to the Chinese speech samples examined for this *setting*, evidence for *palatalization* was found in two of these, *yLong 1* and *wNing 4*. During the examination of *palatalization* it was also discovered that the auditory effect of the final glottalization on the examined syllables - *ke*⁴³ (in *yLong 1*), *he*¹ (in *wNing 4*), *fen*⁴ (in *yMing 3*) and *ke*⁴ (in *yLing 1*)⁴³ - was strikingly soft and gentle. Searching for the origin of this auditory effect, the glottalized phases of these syllables were examined with the help of spectrograms and it was discovered that *glottalization* begins relatively late here, namely 80 ms, 205 ms., 96 ms. and 80 ms. into the respective vowel. By contrast, the examination of final *glottalization* on other, comparable, vowels in the corpus revealed that *glottalization* set in after a mean of 25 ms. The discovered phenomenon was therefore termed *delayed glottalization*⁴⁴. The auditory effect of the four instances of *delayed glottalization* was re-examined and found to be very ‘soft’ and ‘gentle’, as compared to the ‘hard’, ‘rough’ feeling of glottalization affecting longer stretches of vowels. It was therefore concluded that this phenomenon could be made responsible for the ‘soft’ and ‘gentle’ effect of those speech samples in which no evidence for *palatalization* could be found, while in utterances where *palatalization* was present, the auditory effect of *delayed glottalization* supported the ‘gentle’, ‘petite’ auditory effect of *palatalization*.

The presence of *labiodentalized voice* was examined in the Chinese speech samples *yLizhi 4* and *yBing 4*, as their auditory feeling suggested the presence of this *setting*. Unfortunately, however, the instrumental analyses conducted for these speech samples, yielded no conclusive results, so that the presence of *labiodentalized voice* in these utterances could not be definitely clarified⁴⁵. However, since the auditory impression of this *setting* was strong, it was decided to assume that *labiodentalization* (*smiling*) is indeed present in these speech samples. As another factor pointing to the presence of this *setting* here, the dialog from which these utterances were taken, involving You and her woman friend, contains a lot of joking and laughing. The final results for supralaryngeal voice quality are found in Section 4 (line 4) of Table 10.

8.3.3 The final results for overall pitch, tempo and loudness

For the large majority of the 74 speech samples, the auditory and instrumental results for overall pitch/frequency were in good agreement and could be directly adopted as the final results of the phonetic microanalyses⁴⁶. In the few cases where the auditory and instrumental

phonetic microanalyses. They are explained in the introduction to Section 2.

⁴³ As noted in Chapter 3, *glottalization* is expected in syllables bearing a 4th or a 3rd tone, so that the presence of this *setting* in *yLong 1*, *yMing 3* and *yLing 1* is not surprising as such. On the other hand, *glottalization* is not normally found on final 1st tones, so that its presence on *he*¹ (in *wNing 4*) may indeed be regarded as marked. In *fen*⁴ the [n] is absent, but compensated for by the nasalization of the preceding [ə], which is also *glottalized*.

⁴⁴ This delayed onset of *glottalization* in *yLong 1* (on *ke*⁴) is particularly striking as the preceding syllable *ma*³ is also *glottalized*.

⁴⁵ These rather disappointing results for *labiodentalization* are likely to have originated in the fact that the vowels in the examined morphemes

(*lin*² and *shi*⁴) are the front/mid close vowels [i] and [ɨ], which are not only articulated in the same (relatively fronted) area of the oral cavity as *labiodentalization*, but also involve the same spreading lip-movement. As noted in Chapter 5, the presence of supralaryngeal *settings* is difficult to detect in segments which occupy the same area of the articulatory apparatus as the *setting* under investigation.

⁴⁶ In a small number of cases, the auditory results were found to disagree with each other. This was dealt with by adopting the responses given

by the majority of the judges.

results differed in absolute values but agreed with respect to general tendencies, the more differentiated instrumental results were preferred, as the final results were to include subtle variations in pitch⁴⁷.

The auditory and instrumental results for overall tempo were also found to be mostly in good agreement, not only with respect to general tendencies but also regarding absolute values. Small discrepancies were found in a few speech samples. A close examination of the results revealed that the large majority of these differences were linked to the different treatment of pauses and lengthenings in the auditory and instrumental analyses: Whereas in the instrumental analyses of overall tempo - in terms of articulation-rate - calculations had excluded pause phenomena, in the listening tests, by contrast, judgements of overall tempo had been made in terms of speech rate, based on the combined effect of articulation rate and the presence of pauses/lengthenings: While the presence of long phonological pauses/lengthenings may be expected to have incurred auditory judgements of relatively slow overall speech tempo, overly short or missing phonological pauses gave rise to judgements of relatively fast overall tempo⁴⁸. After a re-evaluation of the results, the auditory responses were preferred in utterances containing long pauses/lengthenings with a marked 'slow' auditory feeling⁴⁹. The instrumentally obtained results for overall/mean tempo were, however, preferred if the discrepancies between the judges' responses and the instrumental results did not appear to be linked to the influence of pause phenomena⁵⁰.

Bringing together the auditory and instrumental results for overall loudness was at times more problematic. While in many cases auditory and instrumental results were found to be in agreement, especially with respect to overall tendencies, in a number of speech samples, rather strong differences were found, not only between auditory and instrumental results but also between the auditory judgements made by the judges. As a first step towards solving this problem, whenever this was possible the results indicated by the majority of the judges and confirmed by the instrumental analyses were adopted as the final results. In those cases where this line of action could not be taken, the passages in the original recordings were reanalysed auditorily by the experimenter. The results of these re-analyses which, like the responses of the judges in the listening tests, were based on the original, non-digitized speech material⁵¹, were now found to agree with the auditory responses, which were accordingly adopted as the final results for overall/mean loudness.⁵² The final results for overall loudness, pitch and tempo are given in Section 2 (line 2) of Table 10.

8.3.4 The final results for fluency/non-fluency

According to the definition given in Chapter 4, *fluent* utterances are defined as those containing all - but no more than - the normal phonological pauses/lengthenings as are required for the signaling of (minor and major) intonation-group boundaries. On the other hand, an utterance is referred to as *non-fluent* if it is marked by either the absence of one or more phonological pauses/lengthenings, the presence of marked - i.e. overly long or overly short - phonological pause phenomena, or the presence of non-phonological pauses/lengthenings.

⁴⁷ In a number of cases apparent discrepancies between the auditory and the instrumental results for pitch/frequency are in fact due to the fact that the auditory responses 'low' and 'high' often do not refer to relative pitch-height, as is always the case with the instrumental results, but were used instead with the meanings 'lower' and 'higher than' the pitch-height of previous subsections.

⁴⁸ See Goldman-Eisler 1968.

⁴⁹ These included, for example, *cGans 1*, *cZwieb 3*, *cGans 4*, *sSchoko 1*, *yLa 2*, *yLong 4*, *wLizhi 2* and *wLizhi 3*.

⁵⁰ This was the case in *cVani 2*, *cZwieb 2*, *cMango 2*, *sSchoko 4*, *sSchoko 1*, *sMoon 2*, *sKiwi 3*, *yLing 2*, *yMan 3*, *wMan 2*, *wMan 3* and *wBing 2*.

⁵¹ The digitization of the 74 speech samples necessarily involved amplitude adjustments (See Section 8.2). It must therefore be concluded that the differences in the auditory and instrumental results for loudness/intensity originated in the fact that in contrast to the instrumental analyses, the listening tests were performed on the original, non-digitized, speech material.

⁵² The final results for overall loudness were established in this manner for *cBlum 4*, *cKuli 3*, *cBlum 3*, *cKuli 2*, *cSchoko 2*, *cMango 2*, *cSchoko 2*, *sBlum 3*, *sMango 4*, *sMango 3*, *sBlum 2*, *sKuli 1*, *yLong 4*, *yLing 2*, *yDang 1*, *yLong 2*, *yMan 2*, *yMan 3*, *yLizhi 2*, *wLizhi 4*, *wBing 2*, *wLizhi 3*, *wMan 3*, *wLu 4*, *wLa 4* and *wNing 2*.

Depending on the precise phenomena involved and the auditory feeling present in each case, two different kinds of *non-fluency* are differentiated here: Utterances marked by too short or absent phonological pauses or lengthenings, giving rise to an overly *fast* or *hurried* auditory feeling, are referred to as *hyperfluent*, while those containing overly long phonological pauses or lengthenings, non-phonological pause phenomena or speech discontinuities, whose effect is to create a *slow*, *hesitant* or *disjointed* feeling are referred to as *disfluent*. Turning now to utterance-final slow-down, which also plays a role in speech fluency, insufficient and absent slow-down are defined as features of *hyperfluency*, while overly strong slow-down is regarded as a feature of *disfluency*. Thus, the following phenomena contribute to *fluency/disfluency* and were therefore examined in the phonetic analyses: (1) overly long or short/absent phonological pauses/lengthenings (*disfluency* and *hyperfluency*, respectively); (2) non-phonological pauses of any given length (*disfluency*); (3) speech errors (*disfluency*); and (4) strong or weak/absent, utterance-final slow-down (*disfluency* and *hyperfluency*, respectively). While the presence of phonological and non-phonological pauses/lengthenings of normal and overly long duration were examined in the listening tests as well as in the instrumental analyses, speech errors, marked slow-down rate and overly short/absent pauses/lengthenings were not part of the phonetic listening tests. As determinations of the communicative effect of these phenomena must always take account of the auditory feeling they induce, supplementary assessments of the auditory feeling of the affected utterances were undertaken in all cases and their results were brought together with the instrumental results⁵³.

The following discussion of the results for *fluency*, *disfluency* and *hyperfluency* is divided into four parts: As a first step, we discuss the final results for normal and overly long phonological and non-phonological pauses/lengthenings, established on the basis of both auditory and instrumental results. After this we come to the results for overly short/absent phonological pauses/lengthenings and the presence of speech errors, whose marked auditory feeling of *disfluency* will be included in the discussion. As the third step, we will discuss the final results for degree of utterance-final slow-down, based on the instrumentally obtained results and the auditory feeling of the respective speech samples. Finally, the results for each speech sample mentioned in the discussion will be re-evaluated, yielding their final results, or scores, for fluency and non-fluency, shown in Tables 9.

8.3.4.1 The final results for normal and overly long (non-) phonological pause phenomena⁵⁴

The auditory and instrumental results for normal and overly long (non-) phonological pauses/lengthenings are mostly in agreement. With respect to phonological pauses/lengthenings, however, results differ slightly in 7 cases: In *yLizhi 2* and *yMan 1*, the lengthenings at points (2) and (2), respectively, were not detected in the listening tests, while in *sKuli 1* and *cGans 4* the auditory judgements of long duration relating to the pauses at (1) and (1) were not confirmed instrumentally. In *cGans 1* and *cBlum 1* the phonological pauses at (1) and (1) were established instrumentally as overly long, while in *yLu 3* and *wDan 11* the indicated lengthenings at (1) and (1) could not be confirmed instrumentally. 5 differences were found in the results for non-phonological pause phenomena: In *cKuli 1*, *cMango 2* and *cBlut 4* the non-phonological pauses given at points (3), (2) and (1) were not confirmed instrumentally and in *wMan 3* and *cZwieb 3*, the non-phonological pauses at (2) and (5) were missed in the listening tests⁵⁵. In all these cases, the instrumentally obtained results were taken as final results.

⁵³ The judgement of the auditory feeling of the German speech samples was conducted by the experimenter as a native speaker of German. For the determination of the auditory impression of the Chinese speech samples, the opinion of Chinese native speakers was sought in discussion rounds involving some of the participants in the listening tests.

⁵⁴ Not included in this discussion are differences relating to the absence and possible compensation of phonological pauses, judgements not included in the listening tests. Again, in all cases where increased preboundary lengthening was found instrumentally to compensate for the absence of phonological pauses, the auditory impression of the respective speech samples was also taken into account to decide between *hyperfluency* and *fluency* - the latter in the case of successful, sufficient, compensation and a normal, unmarked auditory feeling.

⁵⁵ A close examination of these speech samples revealed that a large number of the disagreements between the auditory and the instrumental

8.3.4.2 The results for speech errors

A total of 4 speech errors, all followed by a short pause, were found in four utterances: *cZwieb 3*, *sSchoko 1*, *yLa 2* and *wLizhi 3*.

8.3.4.2 The final results for overly short and absent phonological pause phenomena

Overly short phonological pauses, whose presence was determined in the instrumental analyses and confirmed by the auditory feeling of the utterances, are found in *sKuli 4* (2), *yMan 3* (1), *yLizhi 2* (1), *yLong 4* (2), *yMan 1* (2), *wMan 4* (1), *wLizhi 1* (1). In a total of 19 speech samples - *cMango 1* (1), *sBlum 2* (2), *sMoon 1* (1), *sMango 4* (1), *sMango 3* (1,2), *yLu 3* (5), *yLong 3* (1,2), *yMing 3* (2), *yMan 3* (4), *yLu 2* (4), *yLa 2* (1), *yMan 1* (1), *yLong 4* (1,4), *yMing 1* (2), *yDang 1* (3), *wLizhi 2* (1,2), *wLa 4* (2), *wBing 1* (2) and *wNing 2* (1) - one or more phonological pauses are absent and not compensated for by prepausal lengthening. Finally, in *wBing 2*, one prepausal lengthening is absent (2).

8.3.4.3 The final results for utterance-final slow-down

Utterance-final, or preboundary, slow-down was described in Chapter 2 as the tendency of speakers to reduce speech tempo towards the end of intonation-groups in normal speech. Preboundary slow-down is linked to utterance-final syllable-lengthening, found on the last two, maximally three syllables of utterances, therefore determinations of utterance-final slow-down in this study were based on comparisons of syllable/segment duration towards the end of examined speech samples. To minimize the influence of phonologically induced increases of length found on nuclear and stressed/prominent syllables, particularly frequent towards the end of utterances, mean syllable/segment duration was examined over stretches of speech longer than two or three syllables - the ultimate and penultimate subsections of utterances (See Section 8.2.1.2). These calculations were carried out for 8 German and 3 Chinese utterances whose utterance-final auditory feeling was marked, suggesting the presence of irregularities with respect to final slow-down. In order to obtain an understanding of the kind of slow-down rate to be regarded as normal, or unmarked, in the two languages, calculations of final slow-down rate were also carried out for 9 German and Chinese utterances whose auditory feeling was unmarked⁵⁶.

According to Table 7, in the German speech samples slow-down around and slightly above 1.2 were found to be accompanied by a normal, unmarked auditory feeling, indicating that this may be regarded as the normal degree of slow-down in German. On the other hand, slow-down rates ≥ 1.3 were frequently found to incur a feeling of 'slowness' and 'hesitancy', while slow-down rates below 1.2, including negative rates, gave rise to a feeling of 'fastness' and 'hurriedness'. Interestingly, Table 8 suggests a different state of affairs in Chinese: Here, slow-down rates of 1.5 and even 1.6 are obviously unmarked, as they incurred no judgements of marked auditory impression. A slow, hesitant auditory feeling is incurred by a slow-down rate of 1.9, i.e. markedly above 1.6, while a fast, hurried feeling is found in utterances with slow-down rates ≤ 1.3 , including negative rates. It is thus concluded here that these slow-down rates, calculated according to the method applied here, may be regarded as guidelines for the kinds of utterance-final, preboundary, slow-down encountered in German and Chinese and their auditory feeling. Due to their marked auditory feeling, the speech samples *cBlum 4*, *cMoon 1*, *sMoon 1*, *sKuli 1*, *sBlum 1* and *wBing 1* were taken as marked with respect to utter-

results here were obviously due to the presence of sudden, brief tempo variations, a number of which were clearly judged as pauses/lengthenings by the judges (See Goldman-Eisler 1968).

⁵⁶ In the following discussion, the results for the following German speech samples are disregarded: *cMoon 3* ($\times 1.6$), *cBlut 3* ($\times 2.3$), *cBlut 4* ($\times 1.4$), *cGans 4* ($\times 1.9$). A re-examination of these utterances had showed that their ultimate subsections do not lend themselves to calculations of the kind applied here, as they are very short and consist only of two strongly stressed syllables. The result for *sKiwi 2* ($\times 1.04$) is also not included in the discussion, as its ultimate section consists only of two, unstressed, syllables.

ance-final slow-down. *CBlum 4*, the speech sample with a slow, hesitant auditory feeling incurred by overly strong utterance-final slow-down is *disfluent*. On the other hand, *cMoon 1*, *sMoon 1*, *sKuli 1*, *sBlum 1* and *wBing 1*, utterances with a fast, hurried feeling incurred by an overly weak or negative slow-down, are *hyperfluent*⁵⁷. To avoid confusion with other forms of non-fluency, *dis-* and *hyperfluency* incurred by marked utterance-final slow-down are marked with an asterisk.

The German speech samples

Speech samples	Slow-down rate	Auditory impression
cSchoko 3	× (1.2)	normal, unmarked
cKuli3	× (1.21)	normal, unmarked
cZwieb 2	× (1.3)	normal, unmarked
cMango 2	× (1.3)	slow, hesitant (slight)
cSchoko 2	× (1.29)	slow, hesitant (slight)
cBlum 4	× (1.5)	slow, hesitant
cMoon 1	× (1.1)	fast, hurried
cKuli 2	× (-1.47)	fast, hurried (strong)
sGans 1	× (1.25)	normal, unmarked
sMoon 1	× (-1.01)	fast, hurried
sKuli 1	× (-1.09)	fast, hurried
sBlum 1	× (-1.39)	fast, hurried (strong)

The Chinese speech samples

Speech samples	Slow-down rate	Auditory impression
yLong 1	× (1.5)	normal, unmarked
yDang 1	× (1.6)	normal, unmarked
yMan 1	× (1.4)	normal, unmarked
wMan 2	× (1.5)	normal, unmarked
wLa 3	× (1.6)	normal, unmarked
wLizhi 3	× (1.5)	normal, unmarked
wLu 4	× (1.3)	fast, hurried (slight)
wNing 2	× (1.9)	slow, hesitant
wBing 1	× (-1.26)	fast, hurried (strong)

Tables 7 and 8 Utterance-final slow-down rates and their auditory feeling

8.3.4.5 *The final results for fluency and non-fluency*

To obtain the final results for *disfluency* and *hyperfluency* each utterance received a score, according to the degree of markedness in auditory feeling incurred by the phenomena - or items - of *dis-* and *hyperfluency* it contained. Instances of *dis**- and *hyperfluency** originating

⁵⁷ The 'slow' auditory effect of *disfluency* caused by overly strong utterance-final slow-down is due to the fact that this phenomenon reduces, i.e. slows-down, the rate at which the message is delivered by the speaker. The same slowing-down effect is observed in the presence of other phenomena giving rise to *disfluency*, such as speech errors and non-phonological pauses. On the other hand, the auditory effect of overly weak or negative final slow-down (the latter in effect amounting to utterance-final speeding-up) is describable as the impression that the speaker is going too fast and likely to 'crash' any minute. This feeling of 'going too fast' is very similar to that incurred by other phenomena giving rise to *hyperfluency*, such as overly short phonological pauses

in marked slow-down rates received the score ‘½’, while *dis-* and *hyperfluencies* caused by speech errors and irregularities relating to phonological and non-phonological pause phenomena were in most cases given the score of ‘1’. *Disfluency* caused by a very long non-phonological pause received a score of ‘2’, while *hyperfluency* caused by overly short phonological pauses were give the score ‘½’. Finally, the scores of the items in one utterance were added up to obtain the overall score for each speech sample, shown in Table 9⁵⁸. The final scores for *dis-* and *hyperfluency* for each speech sample are also incorporated in Tables 10, where they are found in Section 3 (line 3).

⁵⁸ The numbers indicated in brackets relate to the points at which these phenomena occur in Table B (Section 2) of Appendix III.

Speech sample	<u>Items of disfluency and hyperfluency</u>	score	Degree of dis-/hyperfluency
cBlum 4	A strong slow-down rate of $\times 1.5$ (<i>disfluency</i>)	$\frac{1}{2}^*$	$\frac{1}{2}^*$ <i>disfluency</i>
cMoon 1	1 non-phonological pause preceded by prepausal lengthening (2) (<i>disfluency</i>) A weak slow-down rate of $\times 1.1$. (<i>hyperfluency</i>)	1 $\frac{1}{2}^*$	1 <i>disfluency</i> & $\frac{1}{2}^*$ <i>hyperfluency</i>
cKuli 1	1 non-phonological pause, not preceded by prepausal lengthening (2) (<i>disfluency</i>)	1	1 <i>disfluency</i>
cZwieb 3	1 non-phon. pause with prepausal lengthening (1) (<i>disfluency</i>) 1 non-phon. lengthening (3) (<i>disfluency</i>) 1 non-phonological pause(.) at beginning (5) (<i>disfluency</i>) 1 mistake, preceded by non-phonolog. pause (4) (<i>disfluency</i>)	1 1 1 1	 4 <i>disfluency</i>
cGans 1	1 overly long phonological lengthening (1) (<i>disfluency</i>)	1	1 <i>disfluency</i>
cMango 2	1 non-phonolog. pause preceded by prepausal lengthening (1) (<i>disfluency</i>)	1	1 <i>disfluency</i>
cMoon 3	1 non-phonolog. pause preceded by prepausal lengthening (1) (<i>disfluency</i>)	1	1 <i>disfluency</i>
cBlum 3	1 non-phonolog. pause preceded by prepausal lengthening (1) (<i>disfluency</i>)	1	1 <i>disfluency</i>
cMango 1	1 phonolog. pause absent, not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>
cGans 4	1 non-phonolog. lengthening (2) (<i>disfluency</i>) 1 overly long phonological lengthening (3) (<i>disfluency</i>)	1 1	2 <i>disfluency</i>
cKuli 2	1 non-phonolog. pause preceded by prepausal lengthening (1) (<i>disfluency</i>) A negative slow-down rate of $\times (-1.47)$ (<i>hyperfluency</i>)	1 $\frac{1}{2}^*$	1 <i>disfluency</i> & $\frac{1}{2}^*$ <i>hyperfluency</i>
sKuli 1	A negative slow-rate of $\times (-1.09)$ (<i>hyperfluency</i>)	$\frac{1}{2}^*$	$\frac{1}{2}^*$ <i>hyperfluency</i>
sMango 4	1 phonolog. pause absent and not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>)	1	<i>hyperfluency</i>
sKiwi 2	1 missing phonolog. pause not compensated for by prepausal lengthening (2) (<i>hyperfluency</i>)	1	<i>hyperfluency</i>
sKiwi 1	1 missing phonolog. pause not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>)	1	<i>hyperfluency</i>
sBlum 1	A negative slow-rate of $\times (1.39)$ (<i>hyperfluency</i>)	$\frac{1}{2}^*$	$\frac{1}{2}^*$ <i>hyperfluency</i>
sMango 3	1 phonolog. pause absent, not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>)	1	<i>hyperfluency</i>
sKuli 4	1 phonolog. pause too short, not compensated for by prepausal lengthening (2) (<i>hyperfluency</i>)	1	<i>hyperfluency</i>
sMoon 1	1 non-phonolog. pause absent, is not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>) A negative slow-down rate of $\times (-1.01)$ (<i>hyperfluency</i>)	1 $\frac{1}{2}^*$	1 <i>hyperfluency</i> & $\frac{1}{2}^*$ <i>hyperfluency</i>
sBlum 2	1 non-phonolog. pause preceded by prepausal lengthening (1) (<i>disfluency</i>) 1 phonolog. pause absent, not compensated for by prepausal lengthening (2) (<i>hyperfluency</i>)	1 1	1 <i>disfluency</i> & 1 <i>hyperfluency</i>
sSchoko 1	1 mistake and 1 non-phonolog. pause preceded by prepausal lengthening (1) (<i>disfluency</i>)	1 1	2 <i>disfluency</i>

Table 9 Final results for *disfluency* and *hyperfluency* (German)

Speech samples	2 <u>Items of disfluency and hyperfluency</u>	value	Final results: degree of dis-/ hyperfluency
wBing 2	1 phonolog. lengthening absent (2) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>
wMan 2	2 phonolog. lengthenings absent (3 & 4) (<i>hyperfluency</i>)	2	2 <i>hyperfluency</i>
wLizhi 2	1 phonolog. pause absent (2) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>
wMan 4	1 phonolog. pause is short, not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>)	½	½ <i>hyperfluency</i>
wLizhi 1	1 phonolog. pause overly long w. prepausal lengthening (2) (<i>disfluency</i>) 1 non-phonolog. lengthening (3) (<i>disfluency</i>)	1 1	2 <i>disfluency</i>
wMan 3	1 non-phonolog. pause (2) (<i>disfluency</i>)	1	1 <i>disfluency</i>
wLizhi 3	1 mistake, preceded by syllable lengthening (1) (<i>disfluency</i>)	1	1 <i>disfluency</i>
wLa 4	1 phonolog. pause absent, not compensated for by prepausal lengthening (2) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>
wBing 1	2 non-phonolog. pauses, preceded by prepausal lengthening (3 & 5) (<i>disfluency</i>) 1 phonolog. pause absent, not compensated for by prepausal lengthening (2) (<i>hyperfluency</i>) A negative slow-down rate of $\times (-1.26)$ (<i>hyperfluency</i>)	2 1 ½*	2 <i>disfluency</i> & 1 <i>hyperfluency</i> & ½* <i>hyperfluency</i>
wNing 2	1 phonolog. pause absent, not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>
yLu 3	1 phonolog. pause absent, not compensated for by prepausal lengthening (5) (<i>hyperfluency</i>) 1 non-phonolog. pause with prepausal lengthening (4) (<i>disfluency</i>)	1 1	1 <i>hyperfluency</i> & 1 <i>disfluency</i>
yLong 3	1 phonolog. pause absent, not compensated for by prepausal lengthening (2) (<i>hyperfluency</i>) 1 non-phonolog. pause with prepausal lengthening (3) (<i>disfluency</i>)	1 1	1 <i>hyperfluency</i> & 1 <i>disfluency</i>
yMing 3	1 non-phonolog. pause preceded by prepausal lengthening (3) (<i>disfluency</i>)	1	1 <i>disfluency</i>
yLizhi 4	1 very long non-phonolog. pause with by prepausal lengthening (1) (<i>disfluency</i>)	2	2 <i>disfluency</i>
yMan 3	1 phonolog. pause overly short, not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>) 1 phonolog. pause absent, not compensated for by prepausal lengthening (4) (<i>hyperfluency</i>) 1 long non-phonolog. pause w. strong prepausal lengthening (2) (<i>disfluency</i>) 1 very strong non-phonolog. lengthening (5) (<i>disfluency</i>)	½ 1 3 2	1 ½ <i>hyperfluency</i> & 5 <i>disfluency</i>
Y Lizhi 2	1 phonolog. pause absent, not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>
yLing 2	1 long phonolog. pause (2) (<i>disfluency</i>) 1 non-phonolog. lengthening (1) (<i>disfluency</i>)	1 1	2 <i>disfluency</i>
yLu 2	1 unmarked and 1 long non-phonolog. pause with prepausal lengthening (2 & 5) (<i>disfluency</i>) 1 phonolog. pause absent, not compensated for by prepausal lengthening (4) (<i>hyperfluency</i>)	3 1	3 <i>disfluency</i> & 1 <i>hyperfluency</i>
yLa 2	2 phonolog. pauses absent, not compensated for by prepausal lengthening (1 & 4) (<i>hyperfluency</i>) 1 mistake, followed by a pause (3) and a restart (<i>disfluency</i>)	2 1	2 <i>hyperfluency</i> & 1 <i>disfluency</i>
yLong 4	2 phonolog. pauses absent, not compensated for by prepausal lengthening (1 & 4) (<i>hyperfluency</i>)	2	2 <i>hyperfluency</i>
yMan 1	1 phonolog. pause absent, not compensated for by prepausal lengthening (1) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>
yMing 1	1 phonolog. pause absent, not compensated for by prepausal lengthening (2) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>
yLizhi 1	2 phonolog. pauses overly long (with prepausal lengthening) (1&2) (<i>disfluency</i>)	2	2 <i>disfluency</i>
yDang 1	1 phonolog. pause absent, not compensated for by prepausal lengthening (3) (<i>hyperfluency</i>)	1	1 <i>hyperfluency</i>

Table 9 Final results for *disfluency* and *hyperfluency* (Chinese)

8.3.5 *The final results for intonation*

As the instrumental results for intonation, shown in Table C of Section 2, Appendix III, were in all cases taken as the final results, they need not receive further treatment here. The results for German and Chinese intonation are shown in the 5th section of Tables 10 (lines 5 and 6) and relate to the types of nuclear pitch-patterns, the configurations of the tail-slopes and the types of *head*. Accent-ranges and additional details on tail-configuration are also included, in brackets.

The suprasegmental profiles

cBlum 4 (1.8; +1.0)				
1. Voice Quality: breathy; modal; breathy; modal/lax; modal/lax				
2. Final results for overall loudness, pitch and tempo: soft		mid-low/high	mid-slow	
3. Final results for fluency: ½* disfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Blumen	mid fall (mid)	
		Papierblumen	high fall to mid-high (narrow) levelling-off (gentle)	
	Head(s)	ihr	mid falling (mid)	
cGans 1 (2.6; +1.5)				
1. Voice Quality: modal; modal; modal/lax				
2. Final results for overall loudness, pitch and tempo: mid		v. high/mid/high	slow	
3. Final results for fluency: 1 disfluency				
4. Supralaryngeal V. Q. : palatalization				
5. Intonation	Nuclear pitch-pattern(s)	Gans	high-fall (wide)	gentle
cKuli 1 (2.3; +1.3)				
1. Voice Quality: breathy; breathy; modal; modal/(tense); (modal/lax); modal/lax				
2. Final results for overall loudness, pitch and tempo: mid		mid-high/v. high/mid	mid-slow	
3. Final results for fluency: 1 disfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Kuli	high fall (v. wide)	gentle
		Ach	low level (narrow)	
cZwieb 3 (1.7; +1.1)				
1. Voice Quality: lax; lax; modal/lax; modal/lax; modal/lax; modal; modal/tense				
2. Final results for overall loudness, pitch and tempo: mid		v. high/mid-high/high	mid	
3. Final results for fluency: 4 disfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Sommerzwiebeln	(mid)-high fall (wide)	steep, (rising at end)
		Grün	low rise to v. high (v. wide)	steep
	Head(s)	Möchten	low-mid falling/glissando (mid-wide)	
cMoon 1 (2.5; +1.6)				
1. Voice Quality: modal; modal/lax; lax; (lax); modal				
2. Final results for overall loudness, pitch and tempo: mid		mid/high/ v. high	mid	
3. Final results for fluency: 1 disfluency & ½* hyperfluency				
4. Supralaryngeal V. Q. : palatalization				
5. Intonation	Nuclear pitch-pattern(s)	Haben	rise-fall; low (v. wide)	steep
		Kindergröße	mid rise to high (wide)	steep
cGans 4 (2.9; 1.6)				
1. Voice Quality: breathy; breathy; modal; modal				
2. Final results for overall loudness, pitch and tempo: mid		mid; mid-high/mid-low/low	mid-slow	
3. Final results for fluency: 2 disfluency				
4. Supralaryngeal V. Q. : palatalization				
5. Intonation	Nuclear pitch-pattern(s)	nächstes	high fall to mid-high (narrow)	gentle
		Gans	mid fall (mid)	steep
	Head(s)	dann	high falling (mid-wide)	
cKuli 3 (3.8; -2.2)				
1. Voice Quality: lax; modal; modal; lax, poss.br.				
2. Final results for overall loudness, pitch and tempo: mid-loud		high/mid/mid-low	mid-fast	
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	gerne	mid fall (mid)	gentle
		verkaufe	high falling (narrow)	
	Head(s)			

Table 10-1 Clara

cMango 1 (4.1; -1.9)				
1. Voice Quality: (modal/tense); tense; modal/tense; modal/(tense)				
2. Final results for overall loudness, pitch and tempo:		mid-loud	v. high/mid-low/low	mid-fast
3. Final results for fluency: 1 hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Mangos	high fall (wide)	steep, (flat tail)
		interessieren	high fall (v. wide)	v. steep, (almost flat tail)
	Head(s)	kosten	high level	
		würde	high level	
	Preheads:	Was	low	
		Das	low	
cMango 2 (3.1; -1.7)				
1. Voice Quality: (modal/lax); modal/tense; modal; lax				
2. Final results for overall loudness, pitch and tempo:		mid-soft	mid-high/mid	mid
3. Final results for fluency: 1 disfluency				
4. Supralaryngeal V. Q. : palatalization				
5. Intonation	Nuclear pitch-pattern(s)	Mangos	mid-(high) fall (mid)	steep, (flat tail)
		haben	(mid-high) falling/glissando (narrow)	
cBlut 4 (3.5; -1.7)				
1. Voice Quality: breathy; breathy; breathy; modal; modal				
2. Final results for overall loudness, pitch and tempo:		mid	high/mid/mid-low	mid
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Blutwurst	(mid)-high fall (mid-wide)	steep
		Dann	mid-high rising/glissando (wide)	
cMoon 3 (3.2; +2.0)				
1. Voice Quality: breathy; breathy; modal; lax				
2. Final results for overall loudness, pitch and tempo:		mid	mid/high/v. low	mid-slow
3. Final results for fluency: 1 disfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Moonboots	(mid)-high fall (wide)	steep, (flat tail)
		haben	mid level	
cBlum 3 (3.3; +1.9)				
1. Voice Quality: modal; lax; lax; modal/lax				
2. Final results for overall loudness, pitch and tempo:		mid	mid/v. high/mid-low	mid/slow
3. Final results for fluency: 1 disfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Blumen	low rise to v. high (v. wide)	steep
		Rosen	low rise to v. high (v. wide)	steep
	Head(s)	haben	low level	
cMango 3 (3.1; +1.5)				
1. Voice Quality: (modal/lax); lax; lax; lax; lax				
2. Final results for overall loudness, pitch and tempo:		mid-soft	mid/v. high/low	mid
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Mango	mid-high fall to low (wide)	steep, (flat tail)
		ne	mid rise to v. high (v. wide)	
cSchoko 3 (3.7; +2.3)				
1. Voice Quality: breathy; lax; lax				
2. Final results for overall loudness, pitch and tempo:		mid	v. high/mid-low/low	mid
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Ja	low rise to high (wide)	steep
		Sahnebonbons	low (drop) gentle fall (narrow)	gentle
	Head(s)	Schokolade	high level	
		Preheads:	Schoko-	low

Table 10-1 Clara (cont.)

cBlum 1 (3.8; +1.9)				
1. Voice Quality: lax; modal; modal/lax; modal/tense				
2. Final results for overall loudness, pitch and tempo: mid mid/v. high/mid-low slow				
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Hm	mid rise to v. high (wide)	steep
		Blumen	low rise to v. high (v. wide)	steep
	Head(s)	ein	low level (narrow)	
cSchoko 2 (2.3; +/-1.1)				
1. Voice Quality: modal/lax; modal/lax; modal/(tense); modal/(lax)				
2. Final results for overall loudness, pitch and tempo: mid-soft mid-high/v. high/mid-low mid-fast				
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Kuchen	high fall to low (wide)	steep
	Head(s)	wollte	mid rising/glissando (narrow)	
cKuli 2 (2.8; +/-1.5)				
1. Voice Quality: lax; modal/lax; modal/lax; breathy; modal/lax; modal				
2. Final results for overall loudness, pitch and tempo: mid-soft mid-high/v. high/mid mid/fast				
3. Final results for fluency: 1 disfluency & ½* hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	zwanzig	high fall to low (v. wide)	steep, (flat tail)
	Head(s)	war	mid falling/glissando (mid-wide)	
cBlut 3 (3.5; +/-2.0)				
1. Voice Quality: modal/lax; modal; modal/tense; lax				
2. Final results for overall loudness, pitch and tempo: mid mid-high/high/mid-low fast-slow				
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Blutwurst	(mid)-high fall to low (wide)	steep, (flat tail)
	Head(s)	schauen	(mid-high) falling/glissando (narrow)	
cVani 2 (3.9; -/+2.1)				
1. Voice Quality: modal; modal; modal/tense				
2. Final results for overall loudness, pitch and tempo: mid-soft mid-high/v. high/mid mid/fast				
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Vanillinzucker	high rise-fall (v. wide)	steep, (rising at end)
	Head(s)	Hast	mid-low falling (narrow)	
cZwieb 2 (4.0; +/-2.2)				
1. Voice Quality: modal/(tense); modal; modal/tense; lax				
2. Final results for overall loudness, pitch and tempo: mid mid-high/mid-low/mid fast-mid				
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Karotten	mid-(low) fall (narrow)	gentle
	Head(s)	räumen	high falling/glissando (mid-wide)	

Table 10-1 Clara (cont.)

sKuli 1 (2.4; +1.2)				
1. Voice Quality: (modal/lax); lax; modal; lax; lax, poss. breathy				
2. Final results for overall loudness, pitch and tempo: mid/loud mid/v. high/mid-low mid				
3. Final results for fluency: ½* hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	K <u>u</u> liminen	high fall to low (wide)	steep
		K <u>u</u> li	low fall (narrow)	gentle
	Head(s)	brä <u>u</u> chte	mid-high falling (wide)	
sKiwi 1 (3.5; -2.0)				
1. Voice Quality: modal; modal; modal; modal/lax; lax				
2. Final results for overall loudness, pitch and tempo: mid mid-high/high/mid/mid-low mid				
3. Final results for fluency: 1 hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	drei	mid fall to mid-low (narrow)	gentle
		Kiwi <u>s</u>	mid fall to low (mid-wide)	gentle
	Head(s)	hät <u>t</u> e	high mixed (mid)	
sKiwi 2 (3.6; -1.7)				
1. Voice Quality: modal; modal; modal; (modal/tense); (modal/tense)				
2. Final results for overall loudness, pitch and tempo: mid mid-high/v. high fast				
3. Final results for fluency: 1 hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Hm	low level	
		geh <u>o</u> lt	high rise to v. high (wide)	gentle
	Head(s)	Wie <u>v</u> iel	low rising/glissando (mid-wide)	
sMango 3 (4.4; -2.3)				
1. Voice Quality: modal; modal/tense; modal; modal; modal; modal/lax				
2. Final results for overall loudness, pitch and tempo: mid v. high/high/mid-low/low fast				
3. Final results for fluency: 1 hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	M <u>a</u> ngos	high fall to mid (mid-wide)	gentle, (rising at end)
		J <u>a</u> hreszeit	low fall (drop) (narrow)	(flat tail)
	Head(s)	ist	mid-level	
sSchoko 1 (3.4; -2.1)				
1. Voice Quality: modal; (modal/lax); breathy; lax				
2. Final results for overall loudness, pitch and tempo: mid v. high/mid/mid-low mid				
3. Final results for fluency: 1 disfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Schokol <u>a</u> de	low rise to high (wide)	steep
	Head(s)	k <u>o</u> stet	high falling/glissando (wide)	
sMango 4 (2.8; +1.9)				
1. Voice Quality: lax; modal; modal/lax; lax; lax				
2. Final results for overall loudness, pitch and tempo: mid-loud high/mid-high/mid-low mid				
3. Final results for fluency: 1 hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	schw <u>ie</u> rig	high fall to mid (narrow)	gentle
		Papaya	mid fall to low (mid-wide)	gentle
	Head(s)	M <u>a</u> ngo	high level	
sKiwi 4 (3.1; +2.0)				
1. Voice Quality: modal; modal/lax; breathy; modal				
2. Final results for overall loudness, pitch and tempo: mid high/v. high/mid fast				
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	brauch <u>s</u> t	low rise to v. high (v. wide)	steep
	Head(s)	wie <u>v</u> iel	mid mixed (mid)	

Table 10-2 Sybille

sMoon 1 (3.4; +1.9)				
1. Voice Quality: lax; modal/lax; modal; lax; lax				
2. Final results for overall loudness, pitch and tempo: mid		v. high/mid-high/mid-low		mid-fast
3. Final results for fluency: 1½* hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Gut	mid level	
		ein	high fall to mid (v. wide)	gentle
		dreißig	low fall (narrow)	gentle
	Head(s)	nehme	mid-(high) rising (narrow)	
		Größe	mid-falling (narrow)	
sBlum 3 (3.5; +2.0)				
1. Voice Quality: modal; (modal/lax); lax, poss. breathy; breathy				
2. Final results for overall loudness, pitch and tempo: mid		high/v. high/mid/mid-low		fast
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Blumen	high fall to low (wide)	gentle
		nicht	low fall (narrow)	gentle
sKuli 4 (3.8; +2.2)				
1. Voice Quality: lax; lax; (modal/lax); modal/lax				
2. Final results for overall loudness, pitch and tempo: mid		v. high/mid-high/mid		fast
3. Final results for fluency: 1 hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Hm	mid-high rise to v. high (wide)	steep
		sind	mid fall to low (narrow)	gentle
sKiwi 3 (3.6; +2.3)				
1. Voice Quality: modal; lax; modal/lax				
2. Final results for overall loudness, pitch and tempo: mid		mid-high/v. high/mid-low		mid
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	auch	low fall (drop) (narrow)	(flat tail)
	Head(s)	Kiwis	mid rising to v. high (v. wide)	
sGans 1 (2.9; +/-1.7)				
1. Voice Quality: lax; lax; modal; modal; lax				
2. Final results for overall loudness, pitch and tempo: mid		high/mid-high/v. high		mid-fast
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	haben	mid-(low) rise to v. high (wide)	steep
	Head(s)	morgen	mid-low falling/glissando (wide)	
sBlum 1 (3.1; +/-1.2)				
1. Voice Quality: modal/lax; modal/tense; modal; lax				
2. Final results for overall loudness, pitch and tempo: mid		v. high/mid-high/mid		mid-fast
3. Final results for fluency: ½* hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Blumen	high fall to low (v. wide)	steep, flat tail
	Head(s)	Einen	low level	
sMoon 2 (3.6; +/-2.1)				
1. Voice Quality: modal/lax; modal; modal; modal; breathy				
2. Final results for overall loudness, pitch and tempo: mid		mid-high/high/v. high		mid
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	da	fall-rise (mid-wide)	steep
	Head(s)	Waren	mid-(low) rising/glissando (mid-wide)	

Table 10-2 Sybille (cont.)

sBlum 2 (3.3; +/-2.0)				
1. Voice Quality: modal/lax; modal; modal/tense				
2. Final results for overall loudness, pitch and tempo:		mid-soft	mid-low/high/v. high	mid-fast
3. Final results for fluency: 1 disfluency & 1 hyperfluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Bl <u>u</u> men	low rise to high (v. wide)	steep
		Bl <u>u</u> men	(mid)-high rise to v. high (wide)	steep
sSchoko 3 (3.4; +/-2.1)				
1. Voice Quality: (modal); modal; modal/lax; modal/lax				
2. Final results for overall loudness, pitch and tempo:		mid	high/mid/mid-low	fast
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	Ja	(mid)-low fall-rise (narrow)	steep
		w <u>o</u> llten	mid fall to low (narrow)	gentle
	Head(s)	W <u>e</u> lche	high falling (mid)	
sSchoko 4 (3.6; +/-2.1)				
1. Voice Quality: modal/lax; lax; modal/lax; lax				
2. Final results for overall loudness, pitch and tempo:		mid	mid-high/mid	mid-fast
3. Final results for fluency: fluency				
4. Supralaryngeal V. Q. : —				
5. Intonation	Nuclear pitch-pattern(s)	sein	high fall to mid-high (narrow)	gentle
	Head(s)	W <u>e</u> lche	mixed mid-high (level) (narrow)	

Table 10-2 Sybille (cont.)

yLong 3 (2.4; +1.3)			
1. Voice Quality: modal/lax; modal; modal/lax; lax; modal/lax; modal/tense			
2. Final results for overall loudness, pitch and tempo: mid-soft low/mid-low slow			
3. Final results for fluency: 1 hyperfluency & 1 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 19° Hz Mean tonal band-width: 82 Hz Final tonal behavior: 4 <i>tones</i> (fall); 2.5 <i>tones</i> (rise)			
yLu 3 (2.7; +1.8)			
1. Voice Quality: modal; tense; modal; tense; modal; modal; modal/lax			
2. Final results for overall loudness, pitch and tempo: mid mid-low/low slow			
3. Final results for fluency: 1 hyperfluency & 1 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 40° Hz Mean tonal band-width: 83.5 Hz Final tonal behavior: 1 <i>tone</i> (fall); 1 <i>tone</i> (rise); 1 <i>tone</i> (slope)			
yLizhi 4 (2.7; +1.9)			
1. Voice Quality: modal/lax; tense; tense; modal; tense; tense			
2. Final results for overall loudness, pitch and tempo: mid mid/low slow			
3. Final results for fluency: 2 disfluency			
4. Supralaryngeal V. Q. : labio-dentalization			
5. Intonation Declination: 50° Hz Mean tonal band-width: 106.5 Hz Final tonal behavior: 8 <i>tones</i> (overall fall)			
yMing 3 (3.1; +1.5)			
1. Voice Quality: modal; modal/lax; modal/(tense); modal; modal/(tense)			
2. Final results for overall loudness, pitch and tempo: mid mid-low/low slow			
3. Final results for fluency: 1 disfluency			
4. Supralaryngeal V. Q. : delayed glottalization			
5. Intonation Declination: 22° Hz Mean tonal band-width: 122.6° Hz Final tonal behavior: 5.5 <i>tones</i> (rise); 5.5 <i>tones</i> (slope)			
yMan 2 (4.1; - 2.3)			
1. Voice Quality: modal; modal(tense); modal; modal; modal; modal			
2. Final results for overall loudness, pitch and tempo: mid mid-low/low fast			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 79 Hz Mean tonal band-width: 42.5 Hz Final tonal behavior: 3 <i>tones</i> (fall)			
yLizhi 2 (4.6; - 2.4)			
1. Voice Quality: tense; modal/tense; tense; tense; modal/tense; modal; modal/(lax)			
2. Final results for overall loudness, pitch and tempo: loud mid-low/mid-high/high fast			
3. Final results for fluency: 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 2° Hz (?) Mean tonal band-width: 136.5 Hz Final tonal behavior: 5 <i>tones</i> (overall fall)			
yLong 2 (4.3; - 2.3)			
1. Voice Quality: modal; tense; modal/tense; tense; modal/tense; modal/tense			
2. Final results for overall loudness, pitch and tempo: mid-loud mid-high/low mid-fast			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 72 Hz Mean tonal band-width: 57.3 Hz Final tonal behavior: 3 <i>tones</i> (fall); 1.5 <i>tones</i> (rise)			
yLa 2 (4.4; - 2.5)			
1. Voice Quality: modal/tense; modal/tense; tense; modal/tense; modal; tense; modal; modal; modal/tense; modal/tense			
2. Final results for overall loudness, pitch and tempo: loud-mid low/mid-low/mid-high/high fast			
3. Final results for fluency: 2 hyperfluency & 1 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 54° Hz Mean tonal band-width: 61.5 Hz Final tonal behavior: 2 <i>tones</i> (rise); 3 <i>tones</i> (overall fall)			
yLu 2 (3.6; - 2.1)			
1. Voice Quality: modal/tense; modal; modal/tense; modal/tense; modal/lax			
2. Final results for overall loudness, pitch and tempo: mid mid/mid-low/low mid			
3. Final results for fluency: 3 disfluency & 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 48° Hz Mean tonal band-width: 68 Hz Final tonal behavior: 3.5 <i>tones</i> (overall fall)			

Table 10-3 You

yLing 2 (3.6; - 1.9)			
1. Voice Quality: tense; modal/tense; tense; tense; modal/tense; modal; modal/lax			
2. Final results for overall loudness, pitch and tempo: mid mid-high/low/mid-low mid-fast			
3. Final results for fluency: 2 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 85* Hz Mean tonal band-width: 75.6 Hz Final tonal behavior: 2.5 tones (overall fall)			
yLong 1 (3.6; - 2.1)			
1. Voice Quality: modal; tense; tense; modal; tense; modal; modal/tense			
2. Final results for overall loudness, pitch and tempo: mid mid-high/v. low/low mid			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : palatalization & delayed glottalization			
5. Intonation Declination: 13 Hz Mean tonal band-width: 110.8 Hz Final tonal behavior: 4 tones(rise); 4.5 tones(slope)			
yDang 1 (3.3; - 2.1)			
1. Voice Quality: modal; tense; modal/tense; tense; tense; modal/tense			
2. Final results for overall loudness, pitch and tempo: mid-loud mid-high/mid-low/low/high fast			
3. Final results for fluency: 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 21* Hz Mean tonal band-width: 119.5* Hz Final tonal behavior: 5.5 tones (slope)			
yBing 4 (2.9; +1.9)			
1. Voice Quality: tense; lax; tense; (lax,poss.breathy); (lax,poss.breathy)			
2. Final results for overall loudness, pitch and tempo: mid mid/mid-low/low mid			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : labiodentalized voice			
5. Intonation Declination: 66 Hz Mean tonal band-width: 60.6 Hz Final tonal behavior: 2 tones (overall fall)			
yMan 3 (3.1; +1.9)			
1. Voice Quality: lax; modal/(tense); modal/lax; modal; modal/lax; modal/lax; breathy; lax			
2. Final results for overall loudness, pitch and tempo: mid-soft low/mid-low slow			
3. Final results for fluency: 1 ½ hyperfluency & 5 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 4* Hz (?) Mean tonal band-width: 66.3 Hz Final tonal behavior: 3.5 tones(fall);1 tone(overall fall)			
yLing 1 (3.2; +2.1)			
1. Voice Quality: modal/tense; modal; modal/lax; modal/tense; modal; lax; modal/lax			
2. Final results for overall loudness, pitch and tempo: mid mid-high/low mid			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : delayed glottalization			
5. Intonation Declination: 18* Hz Mean tonal band-width: 97.3 Hz Final tonal behavior: 3.5 tones (fall); 6.5 tones (rise) 6.5 tones (slope)			
yLizhi 1 (3.3; +2.2)			
1. Voice Quality: tense; tense; tense; modal/tense; lax; modal/lax			
2. Final results for overall loudness, pitch and tempo: mid mid-high/low/high mid			
3. Final results for fluency: 2 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 78 Hz Mean tonal band-width: 97 Hz Final tonal behavior: 4.5 tones (fall); 1.5 tones (rise); 1/5 tone (slope)			
yLong 4 (3.3; +2.4)			
1. Voice Quality: tense; tense; modal/tense;tense; (modal/lax); modal/lax; modal/tense			
2. Final results for overall loudness, pitch and tempo: loud mid/mid-low/mid-high mid-fast			
3. Final results for fluency: 2 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 42* Hz Mean tonal band-width: 75.4 Hz Final tonal behavior: 5 tones(fall);1.5 tones(overall rise)			
yMan 1 (3.1; +2.4)			
1. Voice Quality: modal; modal; tense; tense; modal; modal/tense; modal/lax; modal/lax			
2. Final results for overall loudness, pitch and tempo: mid mid/low/mid-high mid-fast			
3. Final results for fluency: 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 53 Hz Mean tonal band-width: 78 Hz Final tonal behavior: 2.5 tones(fall);3 tones(overall rise)			

Table 10-3 You (cont.)

yMing 1 (3.3; +/- 2.0)			
1. Voice Quality: tense; tense; modal; modal/lax; modal/tense; modal; modal; breathy			
2. Final results for overall loudness, pitch and tempo: mid mid-high/mid/mid-low/high/low mid-fast			
3. Final results for fluency: 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation	Declination: 82 Hz	Mean tonal band-width: 97.8 Hz	Final tonal behavior: 2.5 <i>tones</i> (fall);5 <i>tones</i> (overall rise)

Table 10-3 You (cont.)

wNing 4 (1.8; +1.8)			
1. Voice Quality: (modal); (lax); (lax); breathy; breathy; breathy			
2. Final results for overall loudness, pitch and tempo: soft mid/mid-high mid-slow			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : palatalization & delayed glottalization			
5. Intonation Declination: 40 Hz Mean tonal band-width: 77.3 Hz Final tonal behavior: 3.5 <i>tones</i> (fall); 1 <i>tone</i> (rise)			
wLizhi 4 (2.3; +1.4)			
1. Voice Quality: (modal/tense); (modal/lax; poss.breathy); breathy; (modal/lax); (modal/lax); (lax; poss.breathy); breathy			
2. Final results for overall loudness, pitch and tempo: soft mid/mid-high/mid-low mid-slow			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 55* Hz Mean tonal band-width: 46 Hz Final tonal behavior: 2 <i>tones</i> (rise); 2 <i>tones</i> (slope)			
wLa 4 (2.7; +1.9)			
1. Voice Quality: (tense); (modal/lax); breathy; (modal/tense); (modal/lax); breathy; breathy; breathy			
2. Final results for overall loudness, pitch and tempo: mid mid-high/mid/mid-low mid-slow			
3. Final results for fluency: 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 30* Hz Mean tonal band-width: 45.3 Hz Final tonal behavior: 2 <i>tones</i> (fall)			
wBing 3 (3.0; +1.4)			
1. Voice Quality: modal/lax; lax; lax; lax, poss. breathy; breathy; modal/lax; lax			
2. Final results for overall loudness, pitch and tempo: mid-soft mid-high/mid/mid-low mid-fast			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 50 Hz Mean tonal band-width: 64 Hz Final tonal behavior: 1.5 <i>tones</i> (fall)			
wLizhi 3 (2.6; +1.8)			
1. Voice Quality: (modal/lax); modal; modal/lax; modal/lax; (modal/tense); breathy			
2. Final results for overall loudness, pitch and tempo: mid-soft mid-high/mid-low/mid mid			
3. Final results for fluency: 1 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 29* Hz Mean tonal band-width: 94.5 Hz Final tonal behavior: 4 <i>tones</i> (overall fall)			
wBing 1 (3.0; +2.1)			
1. Voice Quality: (modal/tense); (modal/tense); (modal/tense); breathy; (lax); (modal); (modal/tense); (modal/tense)			
2. Final results for overall loudness, pitch and tempo: mid mid-low/low mixed			
3. Final results for fluency: 2 disfluency & 1 hyperfluency & ½*hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 21* Hz Mean tonal band-width: 76.2 Hz Final tonal behavior: 3 <i>tones</i> (rise); 3 <i>tones</i> (fall)			
wLizhi 1 (3.0; +2.4)			
1. Voice Quality: (tense/modal); (modal); (modal/tense); (modal/lax); (modal/lax); (modal); (lax)			
2. Final results for overall loudness, pitch and tempo: mid mid-high/mid-low/mid mid-slow			
3. Final results for fluency: 2 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 16* Hz Mean tonal band-width: 106.9 Hz Final tonal behavior: 2.5 <i>tones</i> (fall)			
wDan 4 (3.3; +2.2)			
1. Voice Quality: (modal/tense); breathy; breathy; breathy; breathy			
2. Final results for overall loudness, pitch and tempo: mid mid/mid-high/mid-low mid			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 70 Hz Mean tonal band-width: 121.6 Hz Final tonal behavior: 5 <i>tones</i> (fall); ½ <i>tone</i> (rise); ¾ <i>tone</i> (slope)			
wMan 4 (3.7; +1.7)			
1. Voice Quality: (modal/tense); (modal/tense); breathy; (lax); (lax); (lax); breathy; breathy; breathy			
2. Final results for overall loudness, pitch and tempo: mid mid-low/mid mid			
3. Final results for fluency: ½ hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 41 Hz Mean tonal band-width: 167.5* Hz Final tonal behavior: 2 <i>tones</i> (rise); 1.5 <i>tones</i> (fall)			

Table 10-4 Wu

wDan 12 (3.4; +2.1)			
1. Voice Quality: (modal/lax); modal; (modal/tense); (modal/tense); breathy; (lax); (lax)			
2. Final results for overall loudness, pitch and tempo: loud mid-high/high mid			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: -23* Hz (?) Mean tonal band-width: 110.6 Hz Final tonal behavior: 4.5 <i>tones</i> (fall)			
wBing 2 (4.5; -2.1)			
1. Voice Quality: (tense); (tense); (tense); (tense); (modal/tense)			
2. Final results for overall loudness, pitch and tempo: loud v. high/mid-high/mid-low fast			
3. Final results for fluency: 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 166 Hz Mean tonal band-width: 143 Hz Final tonal behavior: 7.5 <i>tones</i> (overall fall);2 <i>tones</i> (rise)			
wMan 2 (3.8; -1.9)			
1. Voice Quality: lax; lax, poss.breathy; modal/lax; modal; modal/lax; modal; modal/tense; modal/lax			
2. Final results for overall loudness, pitch and tempo: mid mid-low/low fast			
3. Final results for fluency: 2 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 52 Hz Mean tonal band-width: 91* Hz (?) Final tonal behavior: 9.5 <i>tones</i> (overall fall)			
wLizhi 2 (4.8; -1.8)			
1. Voice Quality: tense; modal; tense; tense; tense; tense; tense			
2. Final results for overall loudness, pitch and tempo: loud high/v. high/mid-high fast			
3. Final results for fluency: 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 79 Hz Mean tonal band-width: 196.8 Hz Final tonal behavior:12 <i>tones</i> (overall fall); 4 <i>tones</i> (rise)			
wLa 3 (3.3; -1.9)			
1. Voice Quality: (modal/tense); (modal/tense); modal/lax; modal/lax			
2. Final results for overall loudness, pitch and tempo: mid v. high/mid-high mid/fast			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 104 Hz Mean tonal band-width: 65 Hz Final tonal behavior: 5.5 <i>tones</i> (rise); 2 <i>tones</i> (fall)			
wDan 11 (3.2; +/- 2.4)			
1. Voice Quality: breathy; (modal/tense); (modal/tense); (modal); breathy			
2. Final results for overall loudness, pitch and tempo: mid low/mid-low mid			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 40* Hz Mean tonal band-width: 187 Hz Final tonal behavior: 10 <i>tones</i> (fall); 9 <i>tones</i> (rise); 2 <i>tones</i> (fall)			
wMan 3 (2.8; +/- 2.1)			
1. Voice Quality: breathy; tense; modal; modal; modal/lax; lax			
2. Final results for overall loudness, pitch and tempo: mid mid-high/mid-low mid-slow			
3. Final results for fluency: 1 disfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 31* Hz Mean tonal band-width: 212.3* Hz Final tonal behavior: 10 <i>tones</i> (fall); 10 <i>tones</i> (rise); 9.5 <i>tones</i> (slope)			
wLu 4 (2.6; +/- 2.0)			
1. Voice Quality: (modal/lax); (modal); (lax); (lax); (lax); breathy			
2. Final results for overall loudness, pitch and tempo: soft mid-low/mid/v. low mid			
3. Final results for fluency: fluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 73 Hz Mean tonal band-width: 114.3 Hz Final tonal behavior: 9 <i>tones</i> (fall); 7 <i>tones</i> (rise); 1.5 <i>tones</i> (fall)			
wNing 2 (2.4; +/- 1.3)			
1. Voice Quality: tense; tense; modal/tense; tense			
2. Final results for overall loudness, pitch and tempo: loud high/mid-high mid			
3. Final results for fluency: 1 hyperfluency			
4. Supralaryngeal V. Q. : —			
5. Intonation Declination: 11 Hz Mean tonal band-width: 199 Hz Final tonal behavior: 8 <i>tones</i> (slope); 5.5 <i>tones</i> (rise-fall)			

Table 10-4 Wu (cont.)

Table 11: Final results of overall pitch, loudness and tempo

No.	Speech sample	overall pitch	overall loudness	overall tempo
1	cBlum 4	mid-low/high	soft	mid-slow
2	cGans 1	v. high/mid/high	mid	slow
3	cKuli 1	mid-high/v. high/mid	mid	mid-slow
4	cZwieb 3	v. high/mid-high/high	mid	mid
5	cMoon 1	mid/high/ v. high	mid	mid
6	cGans 4	mid; mid-high/mid-low/low	mid	mid-slow
7	cKuli 3	high/mid/mid-low	mid-loud	mid-fast
8	cMango 1	v. high/mid-low/low	mid-loud	mid-fast
9	cMango 2	mid-high/mid	mid-soft	mid
10	cBlut 4	high/mid/mid-low	mid	mid
11	cMoon 3	mid/high/v. low	mid	mid-slow
12	cBlum 3	mid/v. high/mid-low	mid	mid-slow
13	cMango 3	mid/v.high/low	mid-soft	mid
14	cSchoko 3	v. high/mid-low/low	mid	mid
15	cBlum 1	mid/v. high/mid-low	mid	slow
16	cSchoko 2	mid-high/v. high/mid-low	mid-soft	mid-fast
17	cKuli 2	mid-high/v. high/mid	mid-soft	mid-fast
18	cBlut 3	mid-high/high/mid-low	mid	fast-slow
19	cVani 2	mid-high/v. high/mid	mid-soft	mid-fast
20	cZwieb 2	mid-high/mid-low/mid	mid	fast-mid

Table 11Clara

No.	Speech sample	overall pitch	overall loudness	overall tempo
1	sKuli 1	mid/v. high/mid-low	mid-loud	mid
2	sKiwi 1	mid-high/high/mid/mid-low	mid	mid
3	sKiwi 2	mid-high/v. high	mid	fast
4	sMango 3	v. high/high/mid-low/low	mid	fast
5	sSchoko 1	v. high/mid/mid-low	mid	mid
6	sMango 4	high/mid-high/mid-low	mid-loud	mid
7	sKiwi 4	high/v.high/mid	mid	fast
8	sMoon 1	v. high/mid-high/mid-low	mid	mid-fast
9	sBlum 3	high/v. high/mid/mid-low	mid	fast
10	sKuli 4	v. high/mid-high/mid	mid	fast
11	sKiwi 3	mid-high/v. high/mid-low	mid	mid
12	sGans 1	high/mid-high/v. high	mid	mid-fast
13	sBlum 1	v. high/mid-high/mid	mid	mid-fast
14	sMoon 2	mid-high/high/v. high	mid	mid
15	sBlum 2	mid-low/high/v. high	mid-soft	mid-fast
16	sSchoko 3	high/mid/mid-low	mid	fast
17	sSchoko 4	mid-high/mid	mid	mid-fast

Table 11 Sybille

No.	Speech sample	overall pitch	overall loudness	overall tempo
1	yLong 3	low/mid-low	mid-soft	slow
2	yLu 3	mid-low/low	mid	slow
3	yLizhi 4	mid/low	mid	slow
4	yMing 3	mid-low/low	mid	slow
5	yMan 2	mid-low/low	mid	fast
6	yLizhi 2	mid-low/mid-high/high	loud	fast
7	yLong 2	mid-high/low	mid-loud	mid-fast
8	yLa 2	low/mid-low/mid-high/high	mid-loud	fast
9	yLu 2	mid/mid-low/low	mid	mid
10	yLing 2	mid-high/low/mid-low	mid	mid-fast
11	yLong 1	mid-high/v. low/low	mid	mid
12	yDang 1	mid-high/mid-low/low/high	mid-loud	fast
13	yBing 4	mid/mid-low/low	mid	mid
14	yMan 3	low/mid-low	mid-soft	slow
15	yLing 1	mid-high/low	mid	mid
16	yLizhi 1	mid-high/low/high	mid	mid
17	yLong 4	mid/mid-low/mid-high	loud	mid-fast
18	yMan 1	mid/low/mid-high	mid	mid-fast
19	yMing 1	mid-high/mid/mid-low/high/low	mid	mid-fast

Table 11 You

No.	Speech sample	overall pitch	overall loudness	overall tempo
1	wNing 4	mid/mid-high	soft	mid-slow
2	wLizhi 4	mid/mid-high/mid-low	soft	mid-slow
3	wLa 4	mid-high/mid/mid-low	mid	mid-slow
4	wBing 3	mid-high/mid/mid-low	mid-soft	mid-fast
5	wLizhi 3	mid-high/mid-low/mid	mid-soft	mid
6	wBing 1	mid-low/low	mid	mixed
7	wLizhi 1	mid-high/mid-low/mid	mid	mid-slow
8	wDan 4	mid/mid-high/mid-low	mid	mid
9	wMan 4	mid-low/mid	mid	mid
10	wDan 12	mid-high/high	loud	mid
11	wBing 2	v. high/mid-high/mid-low	loud	fast
12	wMan 2	mid-low/low	mid	fast
13	wLizhi 2	high/v. high/mid-high	loud	fast
14	wLa 3	v. high/mid-high	mid	mid-fast
15	wDan 11	low/mid-low	mid	mid
15	wMan 3	mid-high/mid-low	mid	mid-slow
17	wLu 4	mid-low/mid/v. low	soft	mid
18	wNing 2	high/mid-high	loud	mid

Table 11Wu

9. The phonetic exponency of the interaction strategies

At this fourth stage of the experimental study, the results of the phonetic microanalysis - the suprasegmental profiles of the 74 utterances - were brought together with the results of the sociolinguistic judgement tests - the classification of these utterances in terms of *deference*, *arrogance*, *support* and *ambiguous valence* - to determine the phonetic exponency of *support*, *deference*, *arrogance* and, as far as possible, *false deference* in German and Chinese¹. To obtain this objective, the suprasegmental profiles of the speech samples in each category were examined carefully to determine their common suprasegmental features which taken together yielded the phonetic exponency of *deference*, *arrogance* and *support* for each language. The suprasegmental profiles in the category of *ambiguous valence* were also examined to verify the obtained results and, in the case of the German utterances, determine the phonetic exponency of *false deference*. The results obtained for each language were then compared with each other in the search for common points and differences - the latter potential sources of miscommunication - across the two languages.

The examination of the phonetic exponency of *deference*, *arrogance* and *support* was carried out separately for each language and consisted of four steps. The manner in which each speaker was found to signal these three strategies was examined first, yielding the results presented in Sections 9.1.1 and 9.1.2 for the German speakers, and in 9.2.1 and 9.2.2 for the Chinese speakers. Following this, the results for the two speakers of each language were brought together, giving a first statement on the phonetic exponency of these strategies in German and Chinese, shown in Sections 9.1.3 and 9.2.3, respectively. As the third step, the contribution of the suprasegmental features was examined in detail for each language on the basis of the obtained results, special attention being devoted to intonation in the hope of clarifying some of the questions which remained open in Chapters 2 and 3. The results of this examination are presented in Sections 9.1.4 and 9.2.4. The results of this third step, yielding more detailed, precise statements on the phonetic exponency of *deference*, *arrogance* and *support* in German and Chinese, are summarized once more in Sections 9.1.5 and 9.2.5. As the fourth and final step, the suprasegmental profiles of the German and Chinese speech samples signaling *ambiguous valence* were examined in order to verify the aforemade statements and determine the phonetic exponency of *false deference* in German. The results of this last examination are summarized in Sections 9.1.6 and 9.2.6. The discussion in Section 9.3 falls into three parts. First, the suprasegmental profiles of *deference*, *support*, *arrogance* and *false deference* in German and Chinese are compared with each other in Section 9.3.1. This is followed in Section 9.3.2 by a comparison of the manner in which the suprasegmental features contribute to the communication of these strategies in the two languages. The results and implications of these comparative examinations are briefly summarized in Section 9.3.3.

9.1 The signaling of *deference*, *arrogance* and *support* in the German speech samples

The following discussion begins with Clara, after which we come to the manner in which Sybille was found to signal *deference*, *arrogance* and *support*.

9.1.1 The signaling of *deference*, *arrogance* and *support* in Clara's speech samples²

As will be recalled from Chapter 7, 15 of Clara's 20 utterances were assigned to the strategies *deference*, *arrogance* or *support*. The following suprasegmental features are most characteristic of Clara's *deferent*, *arrogant* and *supportive* speech samples.

¹ As only two German utterances, *cSchoko* 2 and *cKuli* 2 were found to potentially signal false deference, these examinations began with and focused on the speech samples classified in terms of the other three interaction strategies.

² Speech samples are identified with the labels established according to the method presented in Chapter 7. Their scores for *footing*, strength of position and *valence* are shown between square brackets.

Deference

Clara's 6 deferent speech samples (*cBlum 4*, *cGans 1*, *cKuli 1*, *cZwieb 3*, *cMoon 1* and *cGans 1*) are characterized by (*modal*-)*lax* and (*modal*-)*breathy* voice quality, alternating mostly with *lax*- or *breathiness*. Tenser *settings*, such as *modalness* or *modal-tenseness*, are rare and found exclusively in speech samples with high scores for *footing*/strength of position. This is seen, for instance, in *cGans 1* [2.6; +1.5], which is largely *modal*. *CGans 4*, also strong [2.9; +1.6], is *breathy* at first, *modal* from mid-utterance onwards. *Palatalization* is present in three cases, *cGans 1*, *cGans 4* and *cMoon 1*. Coming to the overall features, Clara's deferent utterances all contain *disfluency*³, which in two cases, *cZwieb 3* and *cGans 4*, is relatively strong. Reduced tempo is present in four of the 6 speech samples. Among these, *cBlum 4*, with very weak scores, [1.8; +1.0], is also marked by soft amplitude. It is also marked by *disfluency**, incurred by overly strong utterance-final slow-down⁴. While overall pitch varies, *cGans 4*, with the strongest overall scores, [2.9; +1.6], is marked by a relatively low pitch. With respect to intonation, accent-range is mostly mid to very wide, a feature found especially in final nuclear pitch-patterns. There are three steep (final) nuclear slopes and three gentle nuclear slopes. Among the latter three, the high fall of *cBlum 4*, holding one of the weakest overall scores [1.8; +1.0], ends with a gently *leveling-off tail* at mid-high pitch. In almost all these utterances, the prenuclear pitch-patterns are low level or gently falling from low or mid, the only exception being *cGans 4*, whose *head* is high falling. Finally, there is a *glissando head* in *cZwieb 3*.

Arrogance

Clara's 4 arrogant speech samples (*cKuli 3*, *cMango 1*, *cMango 2* and *cBlut 4*) share *modalness* and *modal-tenseness*, briefly alternating with *laxness*, as in *cKuli 3*, or *tenseness*, as in *cMango 1*, which holds the strongest score for *footing* [4.1]. *CMango 2*, holding the lowest scores for *footing*/strength of position [3.1; -1.7], is *modal*, mixed with *modal-laxness* and *tenseness*, while *cBlut 4*, the second weakest utterance, is *modal*, alternating with *breathiness*. *Palatalization* is absent in this category, except in *cMango 2*, which is also the only utterance marked by *disfluency* and mid-soft amplitude. With regards to the overall features, *cBlut 4* has mid values for tempo and amplitude, while *cKuli 3* and *cMango 1* have elevated values for both features. On the other hand, *cBlut 4* and *cKuli 3* are *fluent*, while *cMango 1* has slight *hyperfluency*. Pitch varies strongly. Coming to intonation, Clara's arrogant speech samples contain exclusively falling nuclear tones. Accent-range is mid to mid-wide in *cKuli 3*, *cMango 2* and *cBlut 4*, very wide in *cMango 1*, whose two high falling *nuclei* cover a pitch-range of ≤190 Hz. The nuclear slopes are steep and abrupt, *cMango 1* and *cMango 2* ending on *flat tails*. The prenuclear pitch-patterns of *cKuli 3* and *cMango 2* are high and mid-high falling, that of *cBlut 4* mid-high rising, and there is *glissando* in the latter two speech samples. Finally, *cMango 1* has two high-level *heads*, each preceded by a low *prehead*. Its overall contour, which closely resembles that of *stylization*, is referred to here as *quasi-stylization*.

Support

Clara's 5 supportive utterances (*cMoon 3*, *cBlum 3*, *cMango 3*, *cSchoko 3*, and *cBlum 1*) share *lax* and *modal-lax* voice quality, alternating with *breathiness* and *modalness*, only very rarely

³ In the following, the three values of fluency, rendered in cursive script, are defined as: *hyperfluency* (equaling an elevated value for fluency), *disfluency* (a reduced value for fluency) and *fluency* (a mid value for fluency). *Hyper-* and *disfluency* are also described in terms of 'slight', 'strong' and 'very strong', relating to the number of non-fluency items present in the utterances (See Table 9 above).

⁴ It will be recalled from Chapter 8 that *hyperfluency** and *disfluency** are the terms used here to refer to non-fluency resulting from marked utterance-final slow-down. While *disfluency** is incurred by overly strong utterance-final slow-down, *hyperfluency** is caused by an overly weak or negative final slow-down, the latter corresponding to utterance-final speeding-up.

with *tenseness*. While *cBlum 1*, holding high scores for *footing* and strength of position, [3.8; +1.9], has *modalness* alternating with *tenseness*, *cMango 3*, with the lowest overall scores, [3.1; +1.5], is marked by almost exclusive *laxness*. With respect to the overall features, the majority of these speech samples have reduced values for tempo and/or amplitude. Most utterances are *fluent*, only two, *cMoon 3* and *cBlum 3*, containing slight *disfluency*. Coming now to intonation, all the speech samples in this group have wide to very wide accent-ranges. Except for *cMango 3*, which has a final mid-rise, most nuclear patterns are high falls or low rises, two consecutive wide rises being found in *cBlum 1* and *cBlum 3*. All nuclear slopes are steep, with *flat tails* found in *cMango 3*, *cMoon 3* and *cSchoko 3*. In the latter case, the *flat tail* is preceded by a low *prehead*, a high level *head* and a low (*drop*) *nucleus*. As in *cMango 1* the overall shape of this contour resembles that of *stylization*, and is also referred here to as *quasi-stylization*. Finally, the prenuclear pitch-patterns in *cBlum 3*, *cMoon 3* and *cBlum 1* are either low, or mid level.

9.1.2 The signaling of deference, arrogance and support in Sybille's speech samples

Of Sybille's 17 speech samples, 11 were assigned to the categories *deference*, *arrogance* and *support*.

Deference

Only one of Sybille's speech samples, *sKuli 1*, was classified in terms of *deference*. It has *lax* voice quality, alternating with *modal-laxness* and briefly, with *modalness*. It also has *hyperfluency**, caused by an overly weak - in this case negative - utterance-final slow-down rate, mid tempo, elevated amplitude and relatively high overall pitch. With regards to intonation, this utterance contains two falls, the first high and wide with a *flat tail*, the second a low, narrow fall. They are preceded by a mid-high falling *head*.

Arrogance

In Sybille's 4 arrogant utterances (*sKiwi 1*, *sKiwi 2*, *sMango 3* and *sSchoko 1*) *modalness* and *modal-laxness* alternate mostly with *tenseness* and *modal-tenseness*, occasionally with *lax*- or *breathiness*. Brief *laxness* and *breathiness* is found in the two weaker speech samples, *sKiwi 1* [3.5; -2.0] and *sSchoko 1* [3.4; -2.1], while the 2 stronger speech sample, *sKiwi 2* and *sMango 3*, contain *tenseness*. Turning to the overall features, three utterances, *sKiwi 1*, *sKiwi 2* and *sMango 3*, are slightly *hyperfluent*, *sKiwi 2* and *sMango 3*, with high overall scores, [3.6; -1.7] and [4.4; -2.3], also containing elevated tempo. By contrast, *sSchoko 1*, with weak scores, has first degree *disfluency*. While overall pitch varies, *sKiwi 2*, one of the stronger speech samples, is marked by relatively high overall pitch, ranging from very high to mid-high. Coming now to intonation, *sKiwi 2* and *sSchoko 1* have wide rising *nuclei* with steep *slopes* - a high and low rise, respectively - while the final *nuclei* of *sMango 3* and *sKiwi 1* - a low and mid fall - have narrower *slopes*. In *sMango 3*, the final nuclear pattern is a low *drop nucleus* followed by a *flat tail*. *sKiwi 2* and *sSchoko 1* have low rising and high falling *glissando heads*, while *sMango 3* and *sKiwi 1* have a mid level and a high mixed *head*, respectively.

Support

Sybille's 6 supportive speech samples (*sMango 4*, *sKiwi 4*, *sMoon 1*, *sBlum 3*, *sKuli 4* and *sKiwi 3*) share *modal-laxness* alternating with *laxness* and/or *modalness*. *Breathiness* appears briefly in *sBlum 3* and there is *breathiness* and *modal-tenseness* in *sKiwi 4*. *Tenseness* on its own as well as *palatalization* are absent in this category. With respect to the overall features, 4 of these six utterances are marked by relatively fast speech tempo, three contain *hyperfluency*. *Hyperfluency* occurs together with *hyperfluency** in *sMoon 1*. Pitch-height varies, but *sKiwi 4*, holding one of the lowest scores for *footing*, [3.1], is marked by a relatively high overall pitch,

between very high and mid. Coming now to intonation, *sBlum 3* and *sMoon 1* contain two gently falling *nuclei*, the first being high and wide, the second low and narrow. The shapes of these intonation-contours are thus wide, gradually descending slopes. While *sKiwi 3* and *sKiwi 4* also have a wide accent-range, their *tail slopes* are steep. In *sKiwi 3* the low *drop nucleus* is followed by a *flat tail*. Whereas *cKuli 4* contains a mid-falling gentle final *nucleus* and a wide, very steep, mid rising, first *nucleus*, *cMango 4* has two gently descending *nuclear* pitch-patterns, the first beginning at mid-high, the second at mid pitch. With respect to the prenuclear pitch-patterns, there are mid rising *heads* in *sKiwi 3* and *sMoon 1*, a high level *head* in *sMango 4* and a mid mixed *head* in *sKiwi 4*.

9.1.3 Deference, arrogance and support in German

The following is a compilation of the most characteristic suprasegmental features shared by Clara's and Sybille's deferent, arrogant and supportive speech samples, giving a first statement on the phonetic exponency of *deference*, *arrogance* and *support* in German.

9.1.3.1 The deference strategy in German

Clara's and Sybille's 7 deferent speech samples share the pronounced presence of (*modal-*)*lax* and/or (*modal-*)*breathy* voice quality, alternating mostly with *lax-* or *breathiness*. *Modalness* on its own, *modal-tenseness* and *tenseness* are rare and found only in the stronger utterances. Three of Clara's 6 deferent speech samples contain *palatalization*. With respect to the overall features, overall pitch varies, but overall loudness is mostly reduced, tempo is mid to reduced, and all the utterances in this category contain features of non-fluency: There is *hyperfluency** in Sybille's and various, often high, degrees of *disfluency* in all of Clara's utterances.

With respect to intonation, the deferent German speech samples do not appear to have much in common, apart from the fact that the majority have low or mid prenuclear pitch-patterns. Sybille's deferent utterance has a high wide first *nucleus* with a *flat tail*, followed by a narrow low fall. Three of Clara's 6 deferent utterances have wide and steep (final) nuclear slopes⁵, while the 3 more gradual, often less wide, *nuclear* pitch-patterns contain one *leveling-off tail* and one *flat tail*. Clara's *heads* are mostly low level or low/mid falling, only *cGans 4*, her strongest utterance, containing a mid-high falling *head*. Sybille's deferent utterance also has a mid-high falling *head*. *Glissando* is found only once, in *cZwieb 3*.

These results indicate that *deference* in German is signaled by (*modal-*)*lax* and/or (*modal-*)*breathy* voice quality. *Palatalization* is also a frequent feature of this strategy, but *modal voice* and *tenser* phonatory *settings* are found only in stronger utterances, communicating a strong *footing*/strength of position. *Deference* is further communicated by mid to reduced loudness and tempo, and *disfluency* is also frequently encountered. Coming to overall pitch and intonation, the majority of the deferent German utterances were found to share low to mid prenuclear pitch-patterns. Apart from this, however, both features vary quite strongly across the examined utterances so that nothing definite can be said yet regarding their role in this communicative process⁶.

9.1.3.2 The arrogance strategy in German

The 8 arrogant German speech samples share the pronounced presence of *modalness*, mostly alternating with *modal-tenseness* or *tenseness*, less often with *laxness*. *Breathiness* is virtually absent here. A close look at individual scores reveals that utterances containing *tenseness* and/or longer stretches of *modalness* hold higher scores for *footing*/strength of position, while

⁵ In this discussion, the emphasis is on final nuclear pitch-patterns due to their primary role in the signaling of attitudinal messages. (See discussions in Chapter 2)

⁶ More precise statements will follow below, after the examination of the role of the suprasegmental features in these signaling processes, in Section 9.1.5.

utterances containing *laxness* and/or *breathiness* hold lower scores.⁷ Regarding the overall features, the arrogant German utterances are characterized by mid to elevated values for amplitude, tempo and/or *hyperfluency*, the latter found in four speech samples. The strongest utterances, *cMango 1*, *sKiwi 2* and *sMango 3*, have elevated values for all three of these features, while the weakest utterances, *cMango 2* and *sSchoko 1*, contain mid to reduced values for these features. This suggests that a distinction between ‘strongly’ and ‘less strongly’ arrogant utterances is possible: Strongly arrogant utterances share mid, often elevated, values for amplitude and/or tempo as well as, in a number of cases, *hyperfluency*, while less strongly arrogant speech samples have mid, occasionally reduced, values for amplitude and tempo, and mid, at times, reduced values for fluency, i.e. *fluency* and *disfluency*. Overall pitch varies.

Coming now to intonation, a large number of the arrogant German speech samples have relatively wide and steep nuclear pitch-patterns. In three of Clara’s 4 utterances in this category, the falling nuclear *slopes* are steep, with mid to wide accent-ranges. There is *quasi-stylization* in *cMango 1*, the strongest speech sample, and two cases of *flat tails* are found. Among Sybille’s 4 utterances in this category, there are two wide, steep rises. Of her two falls one contains a *drop nucleus* followed by a *flat tail*. The majority of the *heads* are mid or high and there are four cases of *glissando*.

These results indicate that *arrogance* in German is communicated by *modalness* and *modal-tenseness*. *Tenseness* on its own is a feature of strongly arrogant utterances, while *laxness* and, occasionally also *breathiness*, is found in utterances signaling a weaker, inferior, *footing*. *Arrogance* is also signaled with mid to elevated loudness and tempo, accompanied by *fluency* or *hyperfluency*. Reduced values for loudness, tempo and fluency are mostly found in utterances holding low scores for *footing*/strength of position. Again, overall pitch and the intonation features vary quite strongly across the German utterances in this category, so that it is difficult to define their contribution at this point. However, a large number of the arrogant utterances contain wide and steep nuclear pitch-patterns. Most of the prenuclear pitch-patterns are mid or high and there are four cases of *glissando*.

9.1.3.3 The support strategy in German

The 11 supportive utterances in the German corpus clearly share *modal*, *modal-lax* and *lax* voice quality. *Breathiness*, *modal-tenseness* and *tenseness* are rare. With respect to the overall features, the supportive German speech samples seem at first to have little in common: While all five of Clara’s utterances in this category are marked by either mid or mid-soft amplitude and mid or reduced tempo, together with either *fluency* or slight *disfluency*, Sybille’s six supportive speech samples mostly contain mid amplitude and elevated tempo, together with either *fluency* or *hyperfluency*. A close look, however, reveals the common features of mid, in a few cases reduced, amplitude, mid tempo and *fluency*, with Clara tending towards reduced values for tempo and fluency, Sybille towards elevated values for the two features⁸. Overall pitch varies quite considerably across the speech samples of both speakers.

Coming now to intonation, the utterances in this category do not appear to have much in common, apart from the fact that the majority have mid to high prenuclear pitch-patterns. Clara’s utterances typically contain wide to very wide *nuclei* with steep slopes and there are two cases of *flat tails*, one case of *quasi-stylization*. In contrast to this, only two of Sybille’s 6 supportive speech samples have a steep, wide (final) rise or fall, in the latter case a *drop nucleus* with a *flat tail*. Sybille’s other four utterances in this category contain exclusively gradual and gently falling (overall) nuclear pitch-patterns.

⁷ The only exception to this is *cMango 2*, with very weak scores, containing *modalness* and *tenseness*. On the other hand, this speech sample contains *palatalization*, frequent in the deferent utterances, a point to which we will return below.

⁸ These different tendencies on the part of the two German speakers indicates that they possess different communicative, or conversational, styles with respect to the overall features.

These results indicate that *support* in German is signaled by *lax* voice quality, alternating with *modal-laxness*, and by mid loudness, tempo and fluency. While the majority of these utterances have mid to high prenuclear pitch-patterns, otherwise intonation and overall pitch vary quite strongly across the examined utterances, so that their role cannot be ascertained yet.

9.1.4 The role of the suprasegmental phenomena in German

The aim of the following examination is to find out in what way the suprasegmental features contribute to the communication of *deference*, *arrogance* and *support* in German, leading to more precise statements on the phonetic exponency of these strategies. This is achieved by examining the contribution of phonatory voice quality, the overall features, supralaryngeal voice quality and the German forms of intonation to the *valence* and *footing* scores of the German speech samples in each category. It will be recalled from Chapter 1 that positive/negative *valence*, corresponding to the warm/cool or cold undertone of an utterance, enables the listener to distinguish between *support* and *arrogance*, and between *deference* and *false deference*. On the other hand, what is referred to here as the relative *footing* of speech samples, reflected in their scores for footing/strength for position, serves two functions: First, the distinction between strategies with the same *valence* - i.e. *support* and *deference* (positive), and *arrogance* and *false deference* (negative)⁹; And secondly, the distinction between the different degrees of relative footing/strength of position of utterances within one category, corresponding to the relative strength or forcefulness they communicate¹⁰. The examination begins with those features shared most clearly and consistently by the speech samples in each category - phonatory voice quality and overall loudness, tempo and fluency. Following this we will investigate the contribution, or communicative effect, of those features not shared as consistently - overall pitch, the supralaryngeal *settings* and intonation

9.1.4.1 The role of phonatory voice quality

The evidence found in the German suprasegmental profiles indicates that phonatory voice quality is among the features shared most consistently by the speech samples of one category. The arrogant utterances of both speakers share *modalness* and/or *modal-tenseness*, mixing with *tenseness*, while the deferent utterances are marked by (*modal*)-*lax* and (*modal*-) *breathy* voice quality, alternating mostly with *laxness* or *breathiness*, only rarely with *modalness* or *modal-tenseness*. Finally, the speech samples judged as supportive are *modal*, *modal-lax* and/or *lax*, alternating occasionally with *breathiness*, only very rarely with *modal-tenseness*. As the deferent and supportive speech samples share similar combinations of *modal-laxness* and *laxness*, phonatory *settings* with a soft, gentle auditory feeling, and *deference* and *support* share positive *valence*, corresponding to a warm, gentle undertone, this suggests a link between these phonatory *settings* and positive *valence*. Thus, the evidence found in the German suprasegmental profiles suggests that positive *valence* is signaled by combinations of *modal-lax* and *lax* voice quality, alternating with *modalness* or *breathiness*. Coming now to negative *valence*, given the markedly *tenser* voice qualities found in the speech samples of the arrogant category - *modalness* and *modal-tenseness* alternating mostly with *tenseness* - it is concluded that these are linked to the negative *valence* component of this strategy. This fits very well with the colder, harder auditory feeling of *modal-tenseness* and *tenseness*.

⁹ As will be recalled from Chapter 1, weak, inferior *footing* is characteristic of *deference* and *false deference*, while strong, superior *footing* is typically found in *support* and *arrogance*.

¹⁰ As will be recalled from Chapter 2, the 'strength' or 'forcefulness' signaled by an utterance constitutes the attitudinal (paralinguistic) message, as communicated by its suprasegmental phenomena, for example intonation. The terms 'strength' and 'forcefulness' relate to what was referred to in Chapter 2 as the speaker's attitude towards his message. The speaker's interpersonal attitude is circumscribed with the terms 'strong/superior' and 'weak/inferior' *footing*.

As a second point, the evidence in the feature profiles indicates that the presence of certain phonatory *settings* also contribute to the signaling of one or the other interaction strategy. One such a *setting* is *breathiness*, whose presence was found to be particularly strong in the *deference* utterances. Thus, three of Clara's 5 deferent speech samples, *cBlum 4*, *cKuli 1* and *cGans 4*, contain the marked presence of *breathiness*, i.e. over 2 of the examined syllables, a similar presence of *breathiness* being found in only one of her supportive utterances, *cMoon 3*. While Sybille's one deferent utterance was not found to contain *breathiness*, among her 6 supportive speech samples there is only one, *sBlum 3*, which has a comparative presence of *breathiness*, also over 2 syllables. This suggests that *breathiness* in German is used to signal *deference* in German, and to differentiate between *deference* and *support*. A similar role is clearly played by *tenseness*, whose most pronounced presence is found in the utterances of the *arrogance* category, *cKuli 3*, *cMango 4*, *cMango 2*, *sKiwi 2* and *sMango 3*. *Modal/tenseness* and *tenseness* is particularly marked here, covering up to 4 syllables, in *cMango 1*, described in terms of 'strong arrogance' above. This suggests that *tenseness* is used in German to signal *arrogance* or, seen from the listener's point of view, that utterances containing *tenseness* are likely to be interpreted in terms of *arrogance*.

As a third point, there is strong evidence that the phonatory *settings* also contribute to the relative strength, or *footing*, of utterances within one category. Many of the speech samples containing the marked presence of *modal-tenseness* and/or *tenseness* hold relatively high scores for *footing*/strength of position, corresponding to a strong, superior attitude, while utterances with the marked presence of *laxness* and/or *breathiness* generally have lower scores for *footing*/strength of position, corresponding to an inferior *footing*, and attitude. Thus, in Clara's supportive category, *cBlum 1*, the only utterance containing *tenseness*, received the highest score for *footing*, while *cMoon 3*, noted above to contain a strong presence of *breathiness*, received the second weakest score for *footing*. Conversely, among Clara's 4 arrogant utterances, it is *cMango 1*, containing exclusively *modal-tenseness* and *tenseness*, which holds the strongest score for *footing* and the second strongest for strength of position, while *cBlut 4*, containing a blend of *modal voice* and *breathiness*, received the 2nd lowest overall scores for *footing*/strength of position in this category. Similarly, among Sybille's four utterances judged in terms of arrogance, *sSchoko 1*, the only one to contain *breathiness*, has the weakest score for *footing*.

Finally, the German suprasegmental profiles suggest that *modal voice*, is associated with relatively strong *footing*/strength of position, rather than attitudinal neutrality, as was suggested in Chapter 5. In Clara's deferent category, for example, it is *cGans 4*, marked by a strong presence of *modalness* only briefly blending with *breathiness*, which received the strongest overall scores.

To summarize therefore, the suprasegmental profiles indicate four ways in which phonatory voice quality may contribute to the signaling of attitude in German: First, certain (combinations of) phonatory *settings* play an important role in the signaling of *valence*: Positive *valence* is signaled by the *modal-lax* and *lax settings* typical of the deferent and supportive utterances, in which they briefly alternate with *modalness* and *breathiness*, while *modalness* and *modal-tenseness* alternating with *tenseness* - signal the negative *valence* characteristic of the arrogant speech samples. As their second function, the phonatory *settings* also contribute to the relative strength, or *footing*, of utterances signaling one interaction strategy: A strong *footing* is signaled by the pronounced presence of *modal-tenseness* and/or *tenseness*, while a weaker, more inferior *footing* is communicated by the marked presence of *laxness* and/or *breathiness*. As the next point, it was found that *breathiness* is particularly frequent in the deferent utterance, *tenseness* in the arrogant speech samples. This indicates that *breathiness* is used in German to signal *deference* and to differentiate between *deference* and *support*, while *tenseness* is used to communicate arrogance. Finally, the German suprasegmental profiles also

indicate that, rather than signaling a neutral attitude, as was suggested in Chapter 5, *modal voice*, is to be associated with the communication of a strong *footing* in German.

9.1.4.2 The role of overall loudness, tempo and fluency

The examination of the German suprasegmental profiles revealed that none of the overall features taken on its own is shared consistently by the speech samples in any one category. It must be concluded that neither type of interaction strategy nor *valence* or *footing* is signaled by any one these features in isolation. However, a look at Table 12 shows that if one considers overall loudness, tempo and *fluency* together as a group, then combinations of these three features may be said to correlate quite well with the relative strength of the speech samples allocated to one interaction strategy. Thus, the strongest of Clara's supportive speech samples, *cSchoko 3*, contains *fluency*, mid tempo and mid amplitude, while *cBlum 1* and *cMango 3*, with weaker scores for *footing*/strength of position, have *fluency* together with mid or reduced values for amplitude and tempo. Similarly, among Clara's arrogant utterances, the strongest, *cMango 1*, is marked by *hyperfluency* and elevated tempo/amplitude, while the 2nd and 3rd strongest speech samples, *cKuli 3* and *cBlut 4*, have mid fluency and elevated tempo/amplitude and mid tempo/amplitude, respectively. Finally, *cMango 2*, containing the same voice quality *settings* as *cMango 1* but additionally marked by *disfluency*, mid tempo and reduced loudness, holds the lowest overall score for *footing*/strength of position in this group. Similar trends with respect to these features are found in Sybille's supportive and arrogant categories: Of the three arrogant utterances *sKiwi 1*, *sKiwi 2* and *sMango 3*, all containing the same degree of *hyperfluency*, it is *sMango 3* and *sKiwi 2*, marked by fast tempo, which hold the highest scores for *footing*/strength of position, while *sKiwi 1*, containing mid values for amplitude and tempo, holds much lower scores. The weakest utterances here is *sSchoko 1*, containing mid tempo and amplitude together with *disfluency*. Finally, among Sybille's three supportive speech samples containing 1st degree *hyperfluency* - *sMango 4*, *sMoon 1* and *sKuli 4* - it is *sKuli 4*, additionally marked by fast tempo and mid amplitude, which holds the strongest score, the other two containing mid or only slightly elevated values for amplitude and tempo.¹¹ While it is not possible to check the presence of such trends in Sybille's deferent category, which contains only one speech sample, a look at Clara's deferent utterances shows similar patterns here: Among *cBlum 4*, *cGans 1* and *cKuli 1*, all sharing first degree *disfluency*, it is *cBlum 4*, with reduced values for both tempo and amplitude which holds the weakest score, while the other two, holding slightly higher scores, have reduced values for only one of the two features tempo and amplitude. Likewise, among the speech samples *cZwieb 3* and *cMoon 1*, both sharing mid values for tempo and amplitude, it is *cZwieb 3*, with very reduced values for fluency (4th degree *disfluency*) which holds weaker scores, while *cMoon 1*, containing 1st degree of both *hyper-* and *disfluency* received stronger scores for *footing*/strength of position¹². These findings indicate a link between the relative strength, or *footing* of German utterances and their combined values for fluency, amplitude and tempo: Broadly speaking, *hyperfluency* in combination with mid to elevated tempo and/or amplitude signals strong *footing*, while *disfluency* with mid or reduced tempo and/or amplitude communicates a relatively weak *footing*, i.e. inferior attitude. Taking this one step further, if one defines *disfluency* as reduced, *fluency* as mid and *hyperfluency* as elevated fluency, then the following general statement emerges: Among the speech samples

¹¹ In fact, these 3 speech samples contain approximately the same combination of phonatory *settings*, so that their different scores may not be linked to differences in voice quality. Note, however, that *sMoon 1* additionally contains slight *hyperfluency* due to a negative slow-down rate. We will return to this below.

¹² This suggests that the effect of such a strong degree of *disfluency* is in fact 'weaker' than the effect of first degree *dis-* and/or *hyperfluency*. The question as to whether the same, or similar, degrees of *hyper-* and *disfluency* 'cancel each other out', as it were, giving an overall effect of *fluency*, cannot be answered here, as a discussion of such combinations, a large number of which are found in the Chinese corpus would have to take into account the auditory effect of each individual phenomenon.

signaling one and the same one interaction strategy, those containing elevated values for one or more of the three features tempo, amplitude and fluency are often those with stronger scores for *footing*/strength of position, while reduced values for one or more of these features are typical for the weaker utterances belonging to one type of interaction strategy. Turning now to *fluency*, the mid value of the feature fluency, though its precise effect clearly depends on other features present, the suprasegmental profiles indicate that, like *hyperfluency*, *fluency* contributes to the signaling of relative strength, or superior *footing*, especially when in combination with elevated tempo and/or amplitude.

As the last point we now come to the contribution of *hyper**- and *disfluency** brought about by marked utterance-final slow-down¹³. Only one utterance in the German speech corpus, *cBlum 4* (deferent), contains *disfluency**, created by overly strong utterance-final slow-down. *Hyperfluency**, caused by weak or negative slow-down is found in the following 3 speech samples: *sKuli 1* (deferent), *sMoon 1* (supportive) and *cMoon 1* (deferent). Interestingly, 3 of these four speech samples belong to the *deferent* category, while the fourth, *sMoon 1* (*hyperfluency**), holds a relatively weak score for *footing*/strength of position, despite the fact that it also contains combinations of phonatory *settings* associated above with the signaling of relative strength. On the other hand, *cBlum 4* (*disfluency**) is one of the weakest speech samples in its category, although it is the only one not containing *disfluency*. This evidence suggests that both *hyperfluency** and *disfluency** are to be associated with the signaling of relatively weak *footing*, or inferior attitude.

To summarize, the German suprasegmental profiles indicate that combinations of overall tempo, loudness and fluency contribute to the signaling of relative strength, or *footing*, in German. Broadly speaking, elevated values for one or more these features are typically encountered in the stronger speech samples of one category, while reduced values for one or more of these features are found in the weaker utterances. As to *fluency*, the evidence in the German suprasegmental profiles suggests that *fluency* also contributes to the signaling of relative strength. Finally, both *dis**- and *hyperfluency**, caused by marked types of utterance-final slow-down, were found to play a role in the signaling of relative weakness.

9.1.4.3 The role of overall pitch

The evidence in the German suprasegmental profiles rule out a link between overall pitch and type of interaction strategy, as this feature varies, often quite considerably, across utterances. A closer look at individual speech samples, however, shows that an association between pitch and relative strength of utterances is possible to some degree. Thus, among Sybille's arrogant speech samples, *sKiwi 2*, holding the weakest score for strength of position, is marked by a relatively high pitch (ranging between mid-high and very high), while *sMango 3*, the strongest in this category, is marked by a relatively low overall pitch (between mid-high and low)¹⁴. Similarly, among Sybille's supportive utterances, *sKiwi 4*, holding the 2nd lowest score for *footing* despite the presence of *modal-tenseness*, *fluency* and fast tempo¹⁵, also has high overall pitch, between mid and very high, while Sybille's one *deferent* utterance, *sKuli 1*, is also relatively high, between very high and mid-low pitch. Similar trends are found among Clara's speech samples: In her four arrogant utterances the highest overall pitch, ranging from mid-high to mid, is found in *cMango 2*, the weakest in this group, while 4 of her six *deferent* speech samples, *cGans 1*, *cZwieb 3*, *cKuli 1* and *cMoon 1*, are among the utterances with the highest pitch in her corpus (between mid-high and very high). Finally, *cGans 4*, the only

¹³ We will only be discussing speech samples in which *dis**- and *hyperfluency** are found on their own, not in combinations with each other.

¹⁴ Note that *sKiwi 2* and *sMango 3* share the features of fast tempo, mid amplitude and first degree *hyperfluency*, but *sKiwi 2* has a comparatively stronger presence of *modalness* and *modal-tenseness*, which may have led to its relatively high score for *footing* - still much lower, however, than that of *sKiwi 2*.

¹⁵ On the other hand, in *sKiwi 4* *modal-tenseness* and fast tempo may be assumed to have contributed to its relative high score for strength of position.

utterance in this group to descend to very low pitch, is the strongest here¹⁶, while *cSchoko 3*, also marked by very low pitch, holds the strongest score for strength of position among Clara's supportive utterances¹⁷. This suggests that, broadly speaking, high overall pitch in German contributes to the signaling of relative weak *footing*, or inferior attitude, while low pitch is more characteristic of utterances signaling a strong *footing*¹⁸.

9.1.4.4 The role of supralaryngeal voice quality¹⁹

Palatalization was detected in four of Clara's utterances, *cGans 1*, *cMoon 1*, *cGans 4* and *cMango 2*. As the first three are *deferent* speech samples, this suggests an association between *palatalization* and the *deference* strategy. On the other hand, this *setting* is also found in an arrogant utterance, *cMango 2*, suggesting that this association may not be strong as first assumed and that *palatalization* may have other communicative effects. In fact, a closer look at the scores and feature profiles of these 4 utterances indicates that *palatalization* contributes most importantly and directly to the signaling of (relative) *footing*.

We begin with the 3 speech samples in Clara's *deferent* category containing *palatalization*, *cGans 1* [2.6; +1.5], *cMoon 1* [2.5; +1.4] and *cGans 4* [2.9; +1.6]. All three contain two features associated with a weak *footing* above - first degree *disfluency* combined with slow tempo and high overall pitch in the case of *cGans 1*, first degree *disfluency*, high overall pitch and *hyperfluency** in *cMoon 1*, and second degree *disfluency* with reduced tempo in *cGans 4*. On the other hand, all three also contain at least two features associated with a strong *footing*: Mid tempo and amplitude and the pronounced presence of *modalness* in *cGans 1* and *cMoon 1*, and mid amplitude, loudness and fluency and low overall pitch in *cGans 4*. Given the presence of an equal number of 'strong' and 'weak' features, the question arises how these three speech samples came to be allocated to the *deference* category. On the other hand, if one associates *palatalization* with the signaling of 'smallness' and 'non-threat', as suggested in Chapter 5, then the possibility emerges that it was precisely this *setting* which contributed to the judgement of these speech samples in terms of *deference* - by lowering the scores for *footing*/strength of position. Coming now to the effect of *palatalization* in *cMango 2*, much the same may be said here. This utterance is the weakest among Clara's arrogant utterances, with the score of [3.1; -1.7]. While it contains three 'weak' features - reduced amplitude, slight *disfluency* and the highest overall pitch in this group - it also contains a strong presence of *modalness*/*modal-tenseness* comparable to that of *cKuli 3*, holding the 2nd highest score [3.8; -2.2]²⁰. The very low scores of *cMango 2*, despite the marked presence of *modalness* and *modal-tenseness* suggests the same communicative function for *palatalization* - a lowering of scores for *footing*/strength of position - as found in Clara's *deferent* utterances.

In conclusion, the suprasegmental profiles indicate that *palatalization* is associated with the signaling of weakness and inferior *footing* in German. As a second function, the suprasegmental profiles also suggest an additional role in the signaling of *deference*.

9.1.4.5 The role of intonation in German

With respect to the nuclear intonation forms, it was found that nuclear direction, pitch-height/-range and the configuration of tail-slopes vary strongly across the utterances signaling one

¹⁶ This despite *disfluency*, reduced tempo and *palatalization* (whose functions are discussed below).

¹⁷ This utterance also has the 2nd strongest score for *footing* despite the exclusive presence of *laxness* and *breathiness* - unique in this group - and only mid values for fluency, tempo and amplitude. As will be seen below, it is possible that in this case, intonation was involved.

¹⁸ It must be added, however, that this link does not apply in all cases and that the contribution of other suprasegmental features present in an utterance must always also be taken into account.

¹⁹ No *labiodentalization* was detected in the German utterances.

²⁰ As will be seen below, the strong intonation forms present in *cMango 2* are its *glissando head* and *flat tail*, which may be assumed to have also contributed to the negative *valence* judgements this utterance received in the judgement tests.

interaction strategy²¹, leading to the conclusion that type of interaction strategy is not directly communicated by these forms. However, a close look at the suprasegmental profiles shows a link between relative *footing* and type of nuclear pitch-pattern, combining the features of pitch-range and tail-slope configuration. This will be demonstrated in the discussion to follow, taking into account the above established communicative effect of phonatory and supra-laryngeal voice quality, overall loudness, tempo and fluency, and overall pitch.

Regarding the prenuclear pitch-patterns in the German corpus, the majority of the deferent and supportive speech samples have low to mid prenuclear pitch-patterns, while mid to high *heads* are characteristic of the arrogant utterances. This suggests that, broadly speaking, narrow pre-nuclear pitch-patterns communicate positive *valence* - as shared by *deference* and *support*, while the negative undertone characteristic of *arrogance* is communicated by a wide pre-nuclear pitch-range. As a second point, a large number of arrogant utterances were found to contain *glissando heads*, suggesting a role of this pattern in the signaling of negative *valence*, a point which will also be examined below. In order to avoid too much speculation, the following discussion will be limited to those speech samples in which the effect of intonation is most clear, in the main the strongest and weakest utterances of each category.²²

(1) Clara's deferent speech samples

In Clara's deferent category *cBlum 4* [1.8; +1.0] and *cZwieb 3* [1.7; +1.1] hold the weakest scores for footing/strength of position. In both cases, phonatory voice quality is relatively *lax*: There is *modal-laxness* alternating with brief *modalness* and *breathiness* in *cBlum 4*, and *modal-laxness* and *laxness* mixing briefly with *modalness* and *modal-tenseness* in *cZwieb 3*. In addition to this, there is slight *disfluency**, soft amplitude and reduced tempo in *cBlum 4*, marked *disfluency* and high overall pitch in *cZwieb 3*, all phenomena associated with weak *footing* above. Thus, the 'strong' features in these speech samples are: relatively low overall pitch in *cBlum 4* and *fluency* plus mid tempo/amplitude in *cZwieb 3*. Taken together, this adds up to the following combined effects of these (non-intonational) features: *cZwieb 3* has 3 weak features²³ (its phonatory *settings*, very strong *disfluency* and high pitch) and one strong feature (*fluency* and mid tempo/amplitude), while *cBlum 4* has 2 weak features (its phonatory *settings* and *disfluency* plus reduced amplitude and tempo) and one strong feature, its low overall pitch. Coming now to the intonation of these speech samples and their communicative effects, the first *nuclei* are a mid and a (mid-) high fall, respectively, which in *cBlum 4* is followed by a high fall *leveling-off* at mid-high pitch, in *cZwieb 3* by very wide *low rise*. In Chapter 2 the mid fall was associated with the signaling of 'gentleness', while the high *leveling-off* fall in *cBlum 4* strongly resembles the *high narrow fall* of Pike's (1945) framework, associated with the meaning of 'lightness', as opposed to the strength and forcefulness of wide *high falls*. On the other hand, the wide high fall (in *cZwieb 3*) is associated with strength, while Pike (1945) and Tench (1995) associate the (English) *low rise* (also in *cZwieb 3*) with a strong, potentially negative, effect.

Turning to the prenuclear pitch-patterns of the two speech samples, the mid falling *head* of *cBlum 4* was associated in Chapter 2 with attitudinal neutrality, while the English counterpart of the (low) *glissando head* in *cZwieb 3* is associated with 'insistence', as 'forcefulness' in the literature (Tench 1995). If these meanings are applied, then the overall effect of the intonation forms in *cBlum 4* may be described as gentle and weak, that of *cZwieb 3* as signaling a strong *footing* with a potentially negative undertone (*glissando*). A look at the above discussed

²¹ As final level tones are absent in the German speech samples, the contribution of this nuclear pitch-pattern cannot be discussed here.

²² In this examination, the attitudinal functions found in the literature and discussed in Chapter 2 will be applied. In the case of pitch-patterns not covered by the German tonetic literature the functions of their English correlates, as found in the English literature, will be applied tentatively and their validity examined.

²³ In this examination, the effects of overall loudness, tempo and fluency whose contributions were examined above as a group, are taken together. The effect of overall pitch is regarded separately.

effects of the other suprasegmental features suggests that in *cBlum 4* the gentle effect of its intonation forms underscored the effect of the two other weak elements present in this speech sample, making it one of the weakest in Clara's corpus. This may serve to validate the functions suggested here for its intonational pitch-patterns. Coming now to *cZwieb 3*, the strong, potentially negative, effect associated with its intonation forms may be assumed to have reinforced the effect of the strong features in this utterance. Interestingly, however, this speech sample is in fact also one of the weakest in Clara's entire corpus, so that the following two conclusions are possible: Either the combined effect of the assumedly 'strong' intonation forms in this speech sample (the high fall, low wide rise and *glissando head*) is 'weaker' than that of their English counterparts, or the effect of its 3 'weak' elements (*disfluency*, high pitch and the (mid-)high fall) is weak enough to counteract the effect of its two strong features. As the *disfluency* in this speech sample is very strong (4th degree) and its overall pitch is very high, between mid-high and high pitch, it does seem more likely that the weak scores of this utterance were in fact caused by these two features. This supports the meanings applied above to the intonational pitch-patterns of this utterance.

The two strongest utterances in this category are *cGans 4* [2.9;+1.6] and *cGans 1* [2.6;+1.5]. Both contain *palatalization*, *cGans 4* additionally containing *breathiness* alternating with *modal* voice quality, 2nd degree *disfluency* and reduced tempo, but also a rather low overall pitch (between mid and low). *cGans 1*, on the other hand, contains a comparatively stronger presence of *modalness*, 1st degree *disfluency*, slow tempo and a relatively high overall pitch. If one adds up the effects of these features, it emerges that so far both speech samples have approximately the same number of weak and strong features. On the other hand, the score of *cGans 1* is clearly weaker. A look at their intonation shows that *cGans 1* contains a *high falling* gently descending nuclear pitch-pattern, while in *cGans 4* the first *high*, and also narrow, *fall* is followed by a steep mid *fall*. If one associates to these nuclear pitch-patterns the meanings found in Chapter 2 for their English counterparts, then the effect of the *gently* descending *falls* of both utterances may be described as 'weak' or 'gentle', while the *steep* final *fall* in *cGans 4* signals a high degree of strength and forcefulness. Given the higher scores of *cGans 4* as compared with those of *cGans 1*, this suggests that the 'strong', 'forceful' effect of the final steep *fall* of *cGans 1* was very strong indeed, since it was obviously the only feature, apart from its comparatively high overall pitch, responsible for the stronger scores of this utterance. Precisely this attitudinal message was associated with abrupt, steep tail-slopes in English towards the end of Chapter 2, as compared to the gentler, moderate attitudinal effect of gradually descending tail-slopes, and it would seem that these functions also apply to their German counterparts present here.

(2) Clara's arrogant speech samples

The strongest utterance in this category is *cMango 1* [4.1; -1.9]. It contains *modal-tenseness* and *tenseness*, 1st degree *hyperfluency*, elevated loudness and tempo, all these features associated above with a strong and partly negative effect. With respect to intonation, the contour of this speech sample, referred to here as *quasi-stylization*, is very similar to *stylization*, associated in Chapter 2 with a strong, potentially negative effect. The scores of this speech sample, showing that it is the most strongly arrogant, and thus negatively judged, utterance in Clara's entire corpus, may be taken as support for this.

The weakest speech sample in this group is *cMango 2* [3.1; -1.7]. It is marked by the comparatively strong presence of *modal-laxness* and *laxness* and also contains 1st degree *disfluency*, reduced amplitude and mid tempo. In addition to this, there is *palatalization* and a rather high overall pitch. Taken together, these features add up to a high number (four) of weak features, as well as the pronounced presence of *laxness*, with no strong features and no *tense* phonatory settings to account for its classification in terms of *arrogance*. A look at the

intonation of this utterance shows that it contains a mid *fall* with a *flat tail* preceded by a mid-high (falling) *glissando head*. In Chapter 2, the English *glissando head* was associated with strong insistence, authority and forcefulness, while the *flat tail* was associated with strength and potential negativeness. Such strong and potentially negative effects of the *flat tail* and the *glissando head* found in this utterance would certainly have contributed to its classification in terms of *arrogance*, indicating that these functions may be transferred to the *flat tail* and the *glissando head* in German.

(3) Clara's supportive utterances

The two strongest speech samples in this group are *cSchoko 3* [3.7; +2.3] and *cBlum 1* [3.8; +1.9]. *cSchoko 3* is marked by the almost exclusive presence of *laxness*, alternating once with *breathiness*, mid values for fluency, tempo and amplitude and mixed overall pitch. *cBlum 1*, on the other hand, contains a much stronger presence of *modalness* and even brief *tenseness*, also mid fluency but reduced amplitude and tempo and a relatively high overall pitch. Both speech samples therefore seem to contain an equal number of weak and strong features, an assumption which is supported by their scores. Coming to intonation, *cBlum 1* contains two wide, steep rises, from mid to high and from low to high, while in *cSchoko 3* the first low (wide) steep rise is followed by a low (*drop*) *nucleus* with a *flat tail* preceded by a high *head*, whose overall shape is referred to here as *quasi-stylization*. As *quasi-stylization*, like *stylization*, itself, has been associated with a strong, rather negative effect and both speech samples have approximately equal scores, this may be regarded as support for the above suggested association between wide steep rises, particularly in sequence, and a strong, potentially negative attitudinal message.

The weakest speech sample here is *cMango 3* [3.1; +1.5]. Similar to *cSchoko 3*, it is marked by the strong presence of *laxness*, blending only once with *modal-laxness* and has mid values for fluency and tempo, and reduced amplitude. As in *cSchoko 3*, its overall pitch varies. Its two nuclear tones are a mid-high fall with a *flat tail*, followed by a narrow high rise. While the wide fall and *flat tail* of the first *nucleus* has been linked to the signaling of strength, it was seen in Chapter 2 that Pike (1945) associates the English *high rise* with weakness and 'politeness'. Given the fact that *cMango 3* received much lower scores than *sSchoko 3*, whose overall contour (*quasi-stylization*) has been associated with strength and negativeness, and the fact that the first *nucleus* of *cMango 3* is associable with strength, it is plausible to assume that the markedly lower scores of *cMango 3* arose from the presence of the second nuclear pitch-pattern, the narrow high *rise*. This suggests that Pike's meaning of weak, inferior *footing*, signaling a message of gentleness and politeness also applies to the high, narrow rise in German.

(4) Sybille's arrogant speech samples

The strongest utterance in this group is *sMango 3* [4.4; -2.3]. It has mixed overall pitch, *hyperfluency*, fast tempo and mid amplitude, and the marked presence of *modalness*, briefly alternating with *modal-tenseness* and developing into *modal-laxness* at the very end. Although the overall features constitute a strong feature and the effect of *modal-tenseness* and *modalness* are also associable with a strong and negative communicative effect, the question nevertheless arises how this utterance could have received such very strong and negative scores, given the absence of *tenseness* on its own and the presence of *laxness*. A look at the intonation *sMango 3* suggests that the answer lies here: There is a high, yet gradually descending first *fall*, followed by a low *drop nucleus* with a *flat tail*, which is preceded by a mid-level *head*, these adding up to the contour referred to here as *quasi-stylization*. While the effect of the first *nuclear* pitch-pattern was established above to be gentle, that of *quasi-stylization* has already been associated on two occasions with a strong *footing* and a poten-

tially negative undertone. It may therefore be concluded that this second intonation pattern contributed to the strongly negative overall scores of this utterance²⁴.

The weakest utterance in this group is *sSchoko 1* [3.4; -2.1]. It is marked by *modal-laxness*, *laxness* and *breathiness*, mixed overall pitch, *disfluency*, mid amplitude and tempo. While the presence of the overall features, especially *disfluency*, may be assumed to have contributed to the comparatively weak judgement of this utterance as compared to the others in this group, it does seem at first that this speech sample lacks a sufficient number of features signaling negative *valence*, and thus responsible for its classification in terms of *arrogance*. Once again, the answer may lie in the intonation of this speech sample, which consists of a high falling *glissando head*, followed by a wide *low rise*, both of which have been associated not only with a strong *footing*, but also with a potentially negative effect, i.e. negative *valence*. It is very possible that these intonation forms were mainly, perhaps even solely, responsible for the classification of *sSchoko 1* in terms of arrogance.

(5) Sybille's supportive utterances

The two strongest utterances here are *sKuli 4* [3.8; 2.2] and *sKiwi 3* [3.6; 2.3], both containing the same blends of phonatory *settings* - *modal-laxness* mixed with *laxness*. Apart from this, *sKuli 4* contains *hyperfluency* plus elevated tempo and amplitude, while *sKiwi 3* contains mid values for fluency, tempo and loudness. Both speech samples share a relatively high overall pitch. As the overall features of *sKuli 4* have been associated with strong *footing* and the suprasegmental features of the two utterances are otherwise largely identical, *sKuli 4* would be expected to have a stronger score than *sKiwi 3*. This is, however, not the case, both scores adding up to approximately the same strength of *footing*. A look at their intonation shows that *sKuli 4* has a relative wide *mid rise* followed by a high, but gradually falling *nuclear* pitch-pattern, both of which have been associated above with a weak *footing* and gentle attitude. On the other hand, *sKiwi 3* contains a low (*drop*) *nucleus* preceded by a mid-high rising *head*, the auditory effect of this contour again similar to *quasi-stylization*, associated with strength and a potentially negative undertone. A comparison of the communicative effects of these two intonation patterns suggests that the strong, potentially negative effect of *quasi-stylization* in *sKiwi 3* may be equaled with that of the overall features in *sKuli 4*, this leading to the equal scores of the two utterances.

The weakest utterances in this group is *sMango 4* [2.8; +1.9]. It contains the pronounced presence of *laxness*, blending once with *modal-laxness* and once with *modalness*, *hyperfluency* elevated tempo and loudness and mid to mid-high overall pitch. As its combination of phonatory *settings* is comparable to that of *sKuli 4* and its combination of *hyperfluency*, (slightly) elevated tempo and amplitude is also the same, the question arises how *sMango 4* came to receive so much weaker scores than *sKuli 4*, especially as it contains a lower overall pitch. Turning to their intonation, both utterances contain high, gently falling final *nuclear* pitch-patterns, preceded in *sMango 4* by a *high*, relatively *narrow fall*, and in *sKuli 4* by a (mid) rise. The effects of all these three intonation-forms have been described above as gentle, and weak. While the weaker scores of *sMango 4* cannot be definitely clarified, it is possible that the very gentle auditory effect of its (first) high *narrow fall*, already commented on above, contributed to the overall gentler feeling of this utterance.

(6) Sybille's deferent utterance

Sybille's *deferent* utterance, *sKuli 1* [2.4; +1.2], contains *laxness* blending briefly with *modal-laxness* and *modalness*, and there is also evidence of *breathiness*. It also contains *hyperfluency**, elevated amplitude, mid tempo and mixed overall pitch. Its mid-high falling *head* is

²⁴ The rising end of the first nuclear pitch-pattern also has a distinctly negative feeling, evoking an adult reproving or scolding a child.

followed by a steep and wide high fall after which comes a low, but *narrow*, fall. While the communicative effects of the phonatory *settings*, *hyperfluency** and the *narrow* final fall have been described as weak, the effect of the other overall features and that of the wide and steep high (first) fall have been associated a strong *footing*. The low scores of this utterance despite the absence of any marked feature signaling a high degree of weakness suggests that the effect of the final *narrow*, though low, *fall* - described above as weak - overrode that of the first nuclear pitch-pattern, whose communicative effect was described above as strong. This in turn indicates that the communicative effect of intonational pitch-patterns in final position is stronger, in the sense of more immediate and effective on the interpersonal dimension, than non-final pitch-patterns, a possibility already noted above on a number of occasions.

To summarize the above results for intonation, no link was found between nuclear direction and type of interaction strategy, so that a direct association between this form and interpersonal attitude is not possible. This goes counter to statements found in the literature linking nuclear rises to the signaling of inferiority and deference, and nuclear falls to the communication of dominance and superiority²⁵. On the other hand, a number of other intonation forms were found to contribute to the signaling of (relative) *footing* and in some cases also *valence*, or undertone. As a first point, substantial evidence in the German suprasegmental profiles indicates that type of nuclear pitch-pattern, combining nuclear pitch-range and tail-slope configuration plays an important role in the signaling of (relative) *footing*. Thus, *wide* and, particularly, *steep* falls and rises, found frequently among the stronger utterances of one category, communicate a high degree of strength, while *narrow*, i.e. mid or high-mid falling and mid rising, nuclear pitch-patterns, typical of utterances with low scores for *footing*/strength of position, are associable with weak *footing* and a gentle communicative effect, especially when their *tail slopes* descend gradually²⁶. As a second point, the German suprasegmental profiles point to a link between wide, steeply or abruptly falling nuclear pitch-patterns on the one hand and strong *footing* and potentially negative *undertone* on the other, both features combined and particularly pronounced in ‘*drop*’ *nuclei* and *quasi-stylization*. Evidence was also found for an association between the *low wide rise* and a strong, potentially negative effect. By contrast, the attitudinal effect of narrow, gradually descending falls was found to be gentle, particularly when they have additional high pitch. These features are combined in the *high, narrow leveling-off fall* found in the weakest deferent utterance in the German corpus. Coming to the prenuclear pitch patterns, the frequent presence of relatively low, and thus narrow, *heads* in the deferent and supportive German speech samples suggests that narrow pitch-range is associated with the signaling of positive *valence*. Likewise, the mid to high *heads* found to be frequent in arrogant speech samples indicate that wide prenuclear pitch-range plays a role in the signaling of negative *valence*. Finally, strong evidence was found in support of an association between *glissando heads* and negative undertone/*valence*.

9.1.5 Summary: The phonetic exponency of deference, arrogance and support in German

Based on the above results, the phonetic exponency of *deference*, *arrogance* and *support* in German can now be defined more precisely.

Deference

The results of the above discussions indicate that *deference* in German is signaled by (*modal*-) *lax* and/or (*modal*-) *breathy* voice quality, *breathiness* being particularly frequent in highly deferent utterances. By contrast, *tenseness* and long stretches of *modalness* are characteristic

²⁵ As found in Tench 1995. (See Chapter 2 above) It should perhaps be added that on the basis of the evidence obtained here, such statements are not even true for *repeated* falls/rises: As was seen in passing above, not only repeated falls but also repeated *rises*, especially those preceded by *glissando heads* were classified in terms of a strong *footing* and superiority.

²⁶ This finding is in agreement with Uldall's (1960, 1964) results for pitch-range. (See Chapter 2)

of deferent utterances signaling a strong *footing* on the part of the speaker. *Palatalization* is also a characteristic feature of this strategy. Coming to the overall features, *deference* is signaled by mid, frequently reduced, values for loudness, tempo and fluency, with reduced values found especially in highly deferent utterances, *disfluency* also frequently encountered here. Overall pitch varies, but is generally higher in highly deferent utterances, lower in utterances signaling a stronger *footing*. With regards to intonation, the characteristically weak *footing* of deferent utterances is signaled by nuclear pitch-patterns with a narrow range and/or gradually, gently descending tail-slope. The attitudinal message is particularly gentle if these features are combined with high pitch, as in the narrow *leveling-off* high fall. A stronger *footing* is communicated by nuclear pitch-patterns with a wider accent-range. Finally, the relatively narrow *heads* characteristic of the deferent German utterances may be assumed to contribute to their positive undertone.

Arrogance

The above results indicate that *arrogance* in German is communicated by *modalness* and *modal-tenseness*. *Tenseness* is used to communicate strong arrogance, while *laxness* and *breathiness* are found in less strongly arrogant utterances. *Arrogance* is also signaled with mid to elevated loudness, tempo and fluency, reduced values for these features found in utterances signaling a weaker *footing*. Overall pitch is generally low in strongly arrogant utterances, higher in utterances with weaker scores for *footing*. Coming to intonation, the strong *footing* characteristic of strongly arrogant utterances is communicated by steep/abrupt nuclear pitch-patterns with a wide accent-range. The most strongly arrogant speech samples have abruptly falling ‘*drop*’ nuclei or *quasi-stylization*, and the prenuclear pattern of *glissando* is also frequent here. On the other hand, utterances signaling a weaker, less forceful, attitude on the part of the speaker typically contain pitch-patterns with a narrow accent-range. Finally, the relatively wide *heads* characteristic of the arrogant German utterances may be assumed to contribute to their negative undertone.

Support

The above results indicate that *support* in German is communicated by *lax* voice quality which may alternate with *modal-laxness*. Mid values for loudness, tempo and fluency are also characteristic of *supportiveness*. Again, reduced values for these features are frequent in utterances signaling a weak, inferior *footing*, while elevated loudness, tempo and fluency are employed to signal a strong *footing* on the part of the speaker. Again, overall pitch varies, but it is generally higher in weaker utterances, lower in utterances signaling a stronger *footing*. Coming to the role of intonation, strong *footing* is signaled by wide, steep or abrupt, nuclear pitch-patterns, while narrow, gradual pitch-movements are used to communicate a weaker, less forceful attitude. Finally, the relatively narrow *heads* characteristic of the supportive German utterances are also assumed to contribute to their positive undertone.

9.1.6 The German utterances signaling ambiguous valence

A total of 11 German speech samples, *cSchoko* 2, *cKuli* 2, *cBlut* 3, *cVani* 2, *cZwieb* 2, *sGans* 1, *sBlum* 1, *sMoon* 2, *sBlum* 2, *sSchoko* 3 and *sSchoko* 4, were found to be ambiguous with respect to *valence*. The results of their comparative examination are as follows:

As a first result, it was found that the suprasegmental profiles of the German speech samples with ambiguous *valence* widely support the statements made above on the phonetic exponency of *deference*, *support* and *arrogance* and the role of the suprasegmental features in these signaling processes. Strong support in particular was found with respect to what may be described as the *primary* functions of phonatory voice quality - the signaling of undertone/*valence* - and the *primary* functions of the overall features - the communication of the relative

footing/strength of utterances. As a second point, it was found that the combinations of the suprasegmental features present in these speech samples often do *not* conform to the patterns discussed above. In such cases, for example, one finds elevated tempo and mid fluency but reduced amplitude (as in *cSchoko 2*), or elevated tempo combined with reduced loudness, and both *dis-* and *hyperfluency* (as in *sBlum 2*). In other speech samples, such as *sGans 1*, *sMoon 2* and *cVani 2*, there are mid to elevated values for loudness, tempo and fluency together with reduced pitch, while in *cBlut 3*, *cVani 2*, *cSchoko 2* and *sBlum 1*, high overall pitch is found together with *tenseness*. As the overall effect, or message, of these combinations is mixed and thus confusing, it may be assumed that their presence contributed to the ambiguous judgements received by these utterances. As a further point, similar to the second finding, the large majority of these speech samples have a relatively high overall pitch whose gentle effect clashes with very strong and potentially negative intonation features found here: there are 6 cases of *glissando*, 4 *flat tails* and 7 utterances contain one or more steep or abrupt wide falls or rises. On the other hand, none of these ambiguous speech samples contains *palatalization* or the pronounced presence of *tenseness*, suggesting that these voice quality *settings* signal clear, unambiguous messages not likely to be misunderstood - a message of *deference* in the case of *palatalization* and *arrogance* in the case of *tenseness*.

Finally, the two speech samples potentially signaling *false deference*, *cKuli 2* [2.8; +/-1.5] and *cSchoko 2* [2.3; +/-1.1] were both found to contain the pronounced presence of *laxness*, established above to signal positive *valence*. The combined effect, or message, of the overall features in these two speech samples - mid to reduced values for loudness, tempo and fluency and high overall pitch - signals a weak *footing* so that these features are also associable with *deference*. By contrast, the intonation features in both utterances - wide high falls with *flat tails*, additionally preceded by a *glissando head* in *cKuli 2* - are all very strong and potentially negative. These findings indicate that German utterances whose phonatory voice quality and overall features, signaling a warm undertone and inferior *footing*, communicate *deference*, whose strong and potentially negative intonation features, however, go against the signaling of this attitude, are likely to be judged in terms of *false deference*. This in turn indicates that the correct interpretation of a deferent German utterance depends to no small degree on its intonation features, suggesting in effect that the attitudinal role of the intonation forms discussed here is somewhat stronger than was assumed so far.

9.2 The signaling of deference, arrogance and support in the Chinese speech samples

The following discussion begins with speaker Wu, after which we come to manner in which You was found to signal *deference*, *arrogance* and *support*.

9.2.1 The signaling of deference, arrogance and support in Wu's speech samples

As was observed in Chapter 7, fourteen of Wu's 18 speech samples were classified in terms of *deference*, *arrogance* or *support*. The following suprasegmental features were found to be most characteristic of her *deferent*, *arrogant* and *supportive* speech samples.

Deference

Wu's 5 deferent speech samples, *wNing 4*, *wLizhi 4*, *wLa 4*, *wBing 3* and *wLizhi 3*, share the almost exclusive presence of *modal-laxness*, *laxness*, and *breathiness*, with *modalness*, *modal-tenseness* and *tenseness* occurring only rarely and very briefly. *Palatalization* and delayed *glottalization* are found in the weakest utterance, *wNing 4*. With regards to the overall features, the majority of these utterances have reduced amplitude and mid or reduced overall tempo, and mid amplitude and fluency. There is 1st degree *disfluency* in *wLizhi 3* and 1st degree *hyperfluency* in *wLa 4*, the strongest speech samples in this group. Overall pitch is mostly around mid, only *wNing 4* being slightly higher in pitch. Coming to intonation, mean

tonal band-width and final pitch-range vary, and the tonal behavior of the last tone/final particle also varies, between 1 and 2 tones²⁷.

Arrogance

Wu's 4 speech samples in this category, *wBing 2*, *wMan 2*, *wLizhi 2* and *wLa 3*, share the pronounced presence of *modal-tenseness* alternating mostly with *modalness* and/or *tenseness*, only rarely with *lax-* or *modal-laxness*. *WBing 2* and *wLizhi 2*, with very strong scores for *footing/strength* of position, contain the longest stretches of *tenseness* and *modal-tenseness*, while *wLa 3* and *wMan 2*, holding weaker scores, are marked by the pronounced presence of (*modal-*)*laxness*. *Palatalisation* and delayed *glottalization* are absent in this category. With regards to the overall features, overall pitch varies but all these utterances have elevated values for tempo and/or loudness, 3 speech samples additionally containing *hyperfluency*²⁸. Coming now to intonation, mean tonal band-width and final pitch-range vary strongly and the tonal behavior of the last tone/final particle varies between 1 and 4 tones.

Support

Wu's 5 speech samples, *wBing 1*, *wLizhi 1*, *wDan 4*, *wMan 4* and *wDan 12*, are all characterized by the initial phonatory *setting* of *modal-tenseness*, which in all but one of these utterances then develops into *modal-laxness*, *lax-* or *breathiness*. *Tenseness* is absent here as are also *palatalization* and delayed *glottalization*. Overall tempo, loudness and pitch are mostly mid, and three utterances are marked by various degrees of non-fluency: slight *hyperfluency* in *wMan 4*, 2nd degree *disfluency* in *wLizhi 1* and 2nd degree *disfluency* together with 1st degree *hyperfluency* and slight *hyperfluency** in *wBing 1*. With regards to intonation, mean tonal band-width and final pitch-range vary strongly, and the final tonal behavior of the last tone/final particle varies between \emptyset and ≥ 3 tones.

9.2.2 The signaling of deference, arrogance and support in You's speech samples

As discussed in Chapter 7, eighteen of You's 19 speech samples, were assigned to *deference*, *arrogance* or *support*.

Deference

At first sight, You's 4 deferent speech samples, *yLong 3*, *yLu 3*, *yLizhi 4* and *yMing 3*, vary quite strongly with respect to phonatory voice quality. However, a closer look at the supra-segmental profiles and the scores of these utterances shows that they may in fact be divided into two groups with respect to this feature: Thus, *yLong 3* and *yLu 3*, with weaker scores for *footing/strength* of position, are marked by the stronger presence of *modal-laxness* and *breathiness*, alternating only briefly with *modalness* or *modal-tenseness*, while the two stronger utterances, *yLizhi 4* and *yMing 3*, share the more pronounced presence of *modal-tenseness* and *tenseness* and the absence of *breathiness*. In addition, *yLizhi 4* contains *labiodentalization*, while delayed *glottalization* is found in *yMing 3*. With respect to the overall features, overall tempo is slow in all four speech samples, amplitude and pitch are mid in almost all cases, and all four utterances are marked by non-fluency: There is 2nd degree *disfluency* in *yLizhi 4*, 1st degree *disfluency* in *yMing 3*, and 1st degree *dis-* and *hyperfluency* in *yLu 3* and *yLong 3*. Coming now to intonation, mean tonal band-width is mostly wide, the pitch-range of the last tone-group varies, and the tonal behavior of the final tone/particle also varies strongly, between ≥ 1 and ≤ 5.5 tones.

²⁷ The final tonal behavior given here relates to that of the last tones/particles of the Chinese speech samples, found in Table 15. In Chapter 8 the unit of the *Tone (T)* was defined as corresponding to 20 Hz. The results for final tonal behavior will be discussed in detail below.

²⁸ In the following, values for overall features not explicitly mentioned are mid values.

Arrogance

You's 8 arrogant speech samples, *yMan 2*, *yLizhi 2*, *yLong 2*, *yLa 2*, *yLu 2*, *yLing 2*, *yLong 1* and *yDang 1*, all share the strong presence of *tenseness* and/or *modal-tenseness*, alternating with *modalness*. While *breathiness* and *laxness* are absent, there is one occurrence of *modal-laxness*, in *yLu 2*, [3.6; -2.1], one of the weaker utterances in this group, and *palatalization* with delayed *glottalization* is found in *yLong 1*, also one of the weaker utterances here [3.6; -2.1]. The 4 strongest speech samples here - *yMan 2*, *yLizhi 2*, *yLong 2* and *yLa 2* - are all marked by elevated values for loudness and/or tempo, and the two strongest, *yLizhi 2* [4.6; -2.4] and *yLa 2* [4.4;-2.5], have 1st degree *hyperfluency* and 1st degree *dis-* and 2nd degree *hyperfluency*, respectively. In the four utterances with lower scores for *footing/strength* of position, tempo and loudness are mid except in *yDang 1*, slightly louder and faster, and *yLing 2*, where tempo is elevated. Three of the weaker speech samples contain non-fluency: There is 3rd degree *disfluency* and 1st degree *hyperfluency* in *yLu 2*, 2nd degree *disfluency* in *yLing 2* and 1st degree *hyperfluency* in *yDang 1*. Coming to intonation, mean tonal band-width and final pitch-range vary strongly, and final tonal behavior also varies strongly, between 0.5 and ≤ 5.5 tones.

Support

You's 6 speech samples, *yBing 4*, *yMan 3*, *yLing 1*, *yLizhi 1*, *yLong 4* and *yMan 1*, share the initial presence of *modalness*, *tenseness* and/or *modal-tenseness*, developing into the softer settings of *modal-laxness*, *laxness* and *breathiness* around mid-utterance. There is no *palatalization* among these speech samples, but there is delayed *glottalization* in *yLing 1*, and *yBing 4* contains *labiodentalization*. Overall loudness is mid in almost all cases and tempo is mid in half of these speech samples, slow, however, in *yMan 3*, and mid-fast in *yLong 4* and *yMan 1*. Overall pitch varies quite strongly and 4 of these 6 utterances contain various degrees of non-fluency: There is 1st degree *hyperfluency* in *yMan 1*, 2nd degree *hyperfluency* in *yLong 4*, 2nd degree *disfluency* in *yLizhi 1* and very strong *disfluency* accompanied by slight *hyperfluency* in *yMan 3*. Coming to intonation, mean tonal band-width and final pitch-range vary strongly, and final tonal behavior also varies, between 0.25 and ≤ 6.5 tones.

9.2.3 Deference, arrogance and support in Chinese

The following is an examination of the most characteristic suprasegmental features shared by Wu's and You's deferent, arrogant and supportive speech samples, allowing a first statement as to the phonetic exponency of *deference*, *arrogance* and *support* in Chinese.

9.2.3.1 The deference strategy in Chinese

With respect to phonatory voice quality, Wu's 5 deferent utterances share the marked presence of *modal-laxness*, *laxness*, and *breathiness*, alternating only rarely and briefly with *modalness*, *tenseness* and/or *modal-tenseness*. Much the same was found in You's 2 weaker utterances in this category. Her two stronger utterances, however, contain a stronger presence of *modal-tenseness* and *tenseness*. *Palatalization*, *labiodentalization* and delayed *glottalization* are found in three of these 9 speech samples - in *wNing 4*, *yMing 3* and *yLizhi 4*. With regards to the overall features, the deferent Chinese utterances all contain mid to reduced values for tempo and amplitude, reduced values for both features found in 3 speech samples, holding weak scores for *footing/strength* of position. Overall pitch is mid to mid-low and non-fluency is found in a large number of these utterances: There is *disfluency* in one of Wu's and 2 of You's weaker utterances, and *hyperfluency* is found in one of Wu's stronger speech samples. With regards to intonation, mean tonal band-width varies strongly among the deferent utterances of both speakers, and final pitch-range also varies strongly. Finally, the final tonal

behavior of Wu's deferent speech samples is relatively narrow, between ≤ 1 tone and 2 tones, but it is wider in You's *deferent* utterances, varying between ≥ 1 and ≤ 5.5 tones.

These results indicate that *deference* in Chinese is communicated by (*modal-*) *lax*, *lax* and/or *breathy* voice quality. *Palatalization*, delayed *glottalization* and *labiodentalization* are also used in this strategy. *Modalness* and other *tenser* phonatory *settings* are rare and found only in utterances communicating a strong *footing* on the part of the speaker. *Deference* is further communicated by mid to reduced loudness, tempo, and fluency. Regarding overall pitch and the forms of intonation, these features vary strongly across the utterances of the two speakers so that nothing definite can yet be said regarding their contribution to this communicative process.

9.2.3.2 The arrogance strategy in Chinese

With respect to phonatory voice quality, both Wu's and You's *arrogant* speech samples are marked by the strong presence of *tenseness* and *modal-tenseness*, alternating with *modalness* and, in the case of Wu, also with *modal-laxness*. *Palatalisation* and delayed *glottalization* is found only once - in *yLong 1*, one of You's weaker utterances in this category. Regarding the overall features, overall pitch varies but a large number of the arrogant Chinese utterances are marked by elevated values for loudness and/or tempo, elevated values for both features being characteristic of utterances with strong scores for *footing*/strength of position. *Hyperfluency* is also found in a large number of the stronger speech samples, while *disfluency* is typical of utterances with lower scores for *footing*/strength of position. Coming now to intonation, both mean tonal band-width and final tonal behavior vary strongly across the arrogant speech samples of the two Chinese speakers. Finally, the tonal behavior of the final tone/particle varies strongly - between ≥ 1 and 4 tones in Wu's and between 0.5 and ≤ 5.5 tones in You's arrogant utterances.

These results indicate that *arrogance* in Chinese is communicated by *modalness* and *modal-tenseness*. *Tenseness* is found in strongly arrogant utterances, while (*modal-*) *laxness* is more characteristic of utterances communicating a less strong, forceful, attitude. *Arrogance* is also signaled with mid to elevated loudness, tempo and fluency. *Hyperfluency* is found in strongly arrogant utterances, while *disfluency* is more common in utterances holding lower scores for *footing*. Again, overall pitch and the intonation features vary strongly across the Chinese utterances in this category, so that it is difficult to define their role in this process at this point.

9.2.3.3 The support strategy in Chinese

With regards to phonatory voice quality, Wu's and You's *supportive* speech samples are all marked by the presence of *modal-tenseness* at the beginning or in the middle, this *setting* developing into *modalness*, *modal-laxness*, *lax-* or *breathiness* in the course of the utterance. While there is no *tenseness* here, there is also no *palatalization* and delayed *glottalization* is present only in one speech sample, *yLing 1*. With respect to the overall features, overall pitch varies but tempo and loudness are mostly mid. Different kinds of non-fluency are found in a large number of these utterances, *hyperfluency* more characteristic of stronger speech samples, *disfluency* more frequent in utterances with weaker scores for *footing*/strength of position.

With respect to the intonation of these utterances, once again mean tonal band-width and final tonal behavior vary strongly. Final tonal behavior lies between \emptyset and ≥ 3 tones in Wu's and between 0.25 and ≤ 6.5 tones among Wu's speech samples in this category.

These results indicate that *support* in Chinese is communicated by *modal-laxness*, *lax-* and/or *breathiness*. Utterances may also begin with *tenser* phonatory *settings*, which then develop into *laxer*, softer, *settings* by mid-utterance. Support is further signaled by mid tempo and loudness. Strong values for fluency (*hyperfluency*) are found in utterances signaling a strong *footing*, while reduced values for fluency (*disfluency*) are more characteristic of utterances

signaling a weaker, inferior, attitude on the part of the speaker. Once again, overall pitch and intonation vary strongly across the examined Chinese utterances, so that their contribution to this process cannot be defined clearly yet.

9.2.4 The role of the suprasegmental phenomena in Chinese

The aim of the following examination is to determine the manner in which the suprasegmental phenomena contribute to the signaling of *deference*, *arrogance* and *support* in Chinese, allowing more precise statements on the phonetic exponency of these strategies. In analogy to the method applied above to the German speech corpus, this is achieved by examining the role of phonatory voice quality, the overall features, the Chinese forms of intonation and supralaryngeal voice quality to the *valence* and *footing* scores of the Chinese speech samples in each category. Again, the examination begins with the features shared most clearly by the speech samples in each category - phonatory voice quality and overall loudness, tempo and fluency, after which we investigate the contribution, or communicative effect, of those features not shared as consistently - overall pitch, intonation and supralaryngeal voice quality.

9.2.4.1 The role of phonatory voice quality

The evidence in the Chinese suprasegmental profiles indicates that, as in the German corpus, phonatory voice quality is the feature shared most consistently by the Chinese speech samples in one category. The deferent utterances of both Chinese speakers share the marked presence of *modal-laxness*, *laxness* and *breathiness*, alternating only rarely with *modalness*, *modal-tenseness* and *tenseness*. Long stretches of *breathiness* are found in all of Wu's and in the two weakest of You's 4 deferent utterances. The arrogant utterances are characterized by the strong presence of *tenseness* and *modal-tenseness* alternating mostly with *modalness*, only rarely with *modal-laxness*. The supportive speech samples all begin with *modalness*, *tenseness* or *modal-tenseness*, which then develops into the *laxer settings* of *modal-laxness* or *laxness*, and in a number of cases also *breathiness*. Final *breathiness* is found in 3 of Wu's 5 supportive utterances and in the 2 weakest of You's 5 speech samples in this category. As the deferent and supportive speech samples share similar combinations of *modal-laxness*, *laxness* and *breathiness* - all phonatory *settings* with a soft, gentle auditory feeling - and *deference* and *support* share positive *valence*, the above results point to a link between these phonatory *settings* and positive *valence*. The evidence in the Chinese suprasegmental profiles suggests that positive *valence* is signaled by combinations of *modal-lax* and *lax* voice quality, alternating with *modalness* and *breathiness*. Coming now to negative *valence*, the markedly *tenser* voice qualities found in the speech samples of the arrogant category - *modalness* and *modal-tenseness* alternating mostly with *tenseness* - indicates that these are linked to the negative *valence* characteristic of this strategy. Regarding the role of *breathiness*, as this *setting* is found in all of Wu's 4 deferent and 4 of her 5 supportive utterances, and in 2 of You's four deferent and 2 of her five supportive utterances, it must be concluded that, in contrast to German, *breathiness* is a feature of both *support* and *deference* in Chinese and does not serve to differentiate between the two strategies. On the other hand, the pronounced presence of *tenseness* and *modal-tenseness* is certainly the most marked feature of arrogant Chinese utterances: Long stretches of *tenseness*, over two or more syllables, are found in 6 of You's eight and 2 of Wu's four speech samples in this category. It may therefore be concluded that *tenseness* is used in Chinese, as in German, to signal *arrogance*.

As a second point, the Chinese suprasegmental profiles indicate that the phonatory *settings* also contribute to the communication of the relative strength, or *footing*, of utterances. A large number of the speech samples in one category which contain the marked presence of *tense-*

ness and/or *modal-tenseness* hold relatively high scores for *footing*/strength of position, corresponding to a strong, superior *footing*, while the utterances marked by the pronounced presence of *laxness* and/or *breathiness* mostly have lower scores for *footing*/strength of position, corresponding to an inferior *footing*, and attitude. Thus, in Wu's arrogant category, the strongest speech samples, *wLizhi 2* and *wBing 2* are characterized by very long stretches of *tenseness*, extending over minimally four syllables, while the other two speech samples contain merely *modal-tenseness* alternating with *modalness* and *modal-laxness*. Similarly, the 3 weaker utterance in Wu's deferent group, *wNing 4*, *wLa 4* and *wLizhi 3*, are marked by the almost exclusive presence of *lax-* and *breathiness*, while the two stronger speech samples, *wLizhi 4* and *wLa 4*, have a more pronounced presence of *modalness* and *modal-tenseness*. Similar trends are apparent among You's speech samples: Thus, the weakest speech samples in You's *deferent* category, *yLong 3* and *yLu 3*, are marked by the strongest presence of *breathiness* and *laxness*, *modal-tenseness* and *tenseness* being limited to one syllable, while in *yMing 3* and *yLizhi 4*, holding the strongest scores in this group, *modal-laxness* is limited to one syllable and *tenseness* and *modal-tenseness* is found on 2 and 4 syllables, respectively. Among You's 6 *supportive* utterances the three weakest, *yBing 4*, *yMan 3* and *yLing 1*, are marked by the absence of *tenseness* - found only on one syllable in *yBing 4*, while this *setting* extends over a minimum of two syllables in the 3 stronger speech samples, *yLizhi 1*, *yLong 4* and *yMan 1*. This indicates that, broadly speaking, Chinese speech samples signaling a stronger *footing* contain a more pronounced presence of *modal-tenseness* and/or *tenseness*, while utterances with weaker scores for *footing*/strength of position are contain a stronger presence of *laxness* and *breathiness*. Coming now to *modal voice*, the marked presence of this *setting* is found in the following utterances, *yMan 2*, *yLa 2*, *yLing 2*, *yLing 1*, *wLizhi 1* and *wMan 2*. As *wLizhi 1* and *yLing 1* are relatively strong *supportive* utterances and the other 4 utterances all belong to the *arrogance* category, this indicates that, as was also found for German, *modal voice* in Chinese contributes more to the signaling of relative strength of *footing*, than of attitudinal neutrality.

To summarize the above results, the Chinese suprasegmental profiles indicate three ways in which the *phonatory settings* may contribute to the signaling of attitude in Chinese: As their first, possibly primary, function, certain (combinations of) *phonatory settings* play an important role in the signaling of *valence*, a finding also made above for the German speech samples. Thus, positive *valence* is signaled by the (*modal-*)*lax* and *breathy* *phonatory settings* typical of the deferent and supportive Chinese utterances, while negative *valence* is communicated by *tenseness* and *modal-tenseness* whose most pronounced presence is found in the arrogant speech samples of the Chinese corpus. In contrast to German, however, *breathiness* is a feature of both deferent and supportive utterances, indicating that this *setting* does not differentiate between *support* and *deference* in Chinese. On the other hand, it does appear that, as in German, *tenseness* on its own may be regarded as a characteristic feature of *arrogance*. As their second function, the *phonatory settings* also contribute to the relative strength, or *footing*, of utterances signaling one interaction strategy: In Chinese as in German, a strong *footing* on the part of the speaker is signaled by the pronounced presence of *modal-tenseness* and/or *tenseness*, while a weaker, more inferior *footing* is communicated by the marked presence of *laxness* and/or *breathiness*. Finally, the Chinese suprasegmental profiles also indicate that, rather than signaling a neutral attitude, *modal voice* is associated with the communication of a strong *footing*, i.e. a strong, superior attitude, in Chinese, as in German.

9.2.4.2 The role of overall tempo, loudness and fluency

The evidence found in the Chinese suprasegmental profiles indicates that none of the overall features taken on its own is shared consistently by the speech samples in any one category, signaling one type of interaction strategy. It must be concluded that neither type of interaction

strategy nor *valence* nor *footing* is signaled by any one these features in isolation. A look at Table 12, however, shows that, as in German, certain combinations of overall loudness, tempo and *fluency* taken together as a group correlate quite well with the relative strength of speech samples assigned to one category. Thus, in Wu's *arrogant* category the three strongest, *wBing 2*, *wLizhi 2* and *wMan 2*, are marked by elevated values for tempo and fluency and mostly elevated loudness, while *wLa 3*, the weakest in this group, has mid values for loudness and fluency, and only slightly elevated tempo. Similarly, in Wu's *supportive* group, the two strongest speech samples, *wDan 12*, and *wMan 4*, are marked by mid to elevated values for loudness, tempo and fluency, the 2 weakest utterances, *wLizhi 1* and *wBing 1*, have mid values for amplitude, mid/mixed (in *wBing 1*) values for tempo and reduced valued for fluency, which in *wLizhi 1* is strong (2nd degree)²⁹. In Wu's *deferent* group, the two weaker utterances, *wNing 4* and *wLizhi 4*, are marked by reduced values for loudness and tempo, while the in two stronger speech samples, *wLa 4* and *wBing 3*, one of the three overall features is elevated, one is mid and only one has reduced values. Turning to You's speech samples, similar trends are also found here: Thus, *yLong 4* and *yMan 1*, among the three strongest *supportive* speech samples, are marked by mid to elevated loudness and tempo, and 1st and 2nd degree *hyperfluency*, respectively, while reduced values for tempo and loudness and strong *hyperfluency* (5th degree) is found in *yMan 3*, one of the weakest speech samples in this category. Similarly, the 2 strongest of You's *arrogant* speech samples, *yLizhi 2* and *yLa 2*, contain elevated values for tempo and amplitude, as well as 1st and 2nd degree *hyperfluency*, while the two weakest speech samples here, *yLong 2* and *yMan 2*, also marked by elevated tempo and mid/elevated amplitude, contain not *hyperfluency* but *fluency*, i.e. mid values for fluency. These data indicate that, similar to what was found for German, the relative strength, or *footing*, of Chinese utterances is communicated by their combined values for fluency, amplitude and tempo: Broadly speaking, utterances within one category holding high scores for *footing*/strength of position typically contain mid to elevated valued for loud-ness, tempo and fluency, while mid to reduced values for these features are found most often in utterances signaling a weaker, more inferior *footing*³⁰.

Turning now to the communicative role of *hyperfluency**, caused by a marked utterance-final slow-down³¹, this feature was found in only one speech sample, *wBing 1*, holding the weakest score in Wu's *supportive* group. Though the effect of this feature is a difficult to estimate, since it occurs together with 2nd degree *disfluency* and 1st degree *hyperfluency*, it is rather striking that *wBing 1* received the lowest scores in this category despite its very pronounced presence of *modal-tenseness* (over 5 syllables) and mid/mixed values for loudness and tempo. The lack of further weak features in this utterance suggests that the effect of *disfluency** was indeed weak here and that it supported the weak effect of 2nd degree *disfluency*, contributing to the surprisingly weak score of this speech sample³².

²⁹ In *wBing 1*, 2nd degree *disfluency* is mixed with 1st degree *hyperfluency* and slight *hyperfluency**, so that evidence for the proposed link is not quite as clear as in *wLizhi 1*.

³⁰ With respect to overall loudness, it should perhaps be noted in passing that the data in Table 12 also provide support for a further, specifically Chinese, link between this feature and interpersonal attitude, to which attention has been called a number of times in this work - between elevated loudness and the signaling of a strong, potentially negative attitude. Thus, 4 out of Wu's 5 speech samples containing reduced loudness were classified in terms of *deference*, and of You's two utterances containing reduced loudness, one was assigned to the *deference*, the other to the *support* strategy. On the other hand, of the 5 speech samples in You's corpus containing elevated loudness, as many as 4 were judged in terms of *arrogance*, and of Wu's 4 utterances with elevated loudness 2 were also classified in terms of *arrogance*. By contrast, only 2 of Clara's 6 utterances with reduced values for loudness were classified in terms of *deference* or *support*, the others being judged as *arrogant* or *ambiguous*, and Sybille's one utterance with reduced loudness was also judged to be *ambiguous*. And while Clara's 2 utterances with elevated loudness were both assigned to the *arrogance* category, of the 2 produced by Sybille one was judged in terms of *deference*, the other in terms of *support*. This indicates that, to a much higher degree than in German, utterances with elevated loudness tend to be judged in Chinese in terms of a strong, potentially *arrogant* attitude, while utterances containing reduced values for loudness are mostly judged in terms of *deference*.

³¹ In *wBing 1*, final slow-down is absent. Other cases of marked utterance-final slow-down were not found in the Chinese corpus.

³² Specific communicative functions of the kind discussed for the German speech samples could not be found for *fluency*, the mid value for overall fluency.

To summarize, the Chinese suprasegmental profiles indicate that combinations of overall tempo, loudness and fluency contribute to the signaling of relative strength, or *footing*, in Chinese. As was also found for German, elevated values for one or more of these features are typically found in the stronger speech samples of one category, while reduced values for one or more of these features are encountered utterances signaling a weaker footing³³. The Chinese suprasegmental profiles also indicate that *hyperfluency**, caused by marked utterance-final slow-down, contributes to the signaling of relative weakness.

9.2.4.3 The role of overall pitch

In the examination of the Chinese suprasegmental profiles, overall pitch, described in terms of the *bottom-line* of the Chinese utterances, was found to vary quite strongly. However, a close look at the pitch of individual utterances, as shown in Table 11 shows that a link between this feature and interpersonal attitude does seem possible. Thus, all of You's deferent and supportive speech samples are marked by a relatively low overall pitch (between low and mid-high) while 3 of her 8 arrogant utterances, in particular her most strongly arrogant ones, *yLizhi 2*, *yLa 2* and *yDang 1*, are marked by elevated values for this feature (between mid-low and high). Similarly, all but one of Wu's 5 deferent utterances and her 5 supportive utterances have reduced overall pitch, while elevated values for this feature are found in 3 of her 4 arrogant speech samples. As *support* and *deference* share the component of positive *valence* while *arrogance* is characterized by negative *valence*, this suggests that a role of overall pitch in Chinese in the signaling of *valence*: High overall pitch contributes to the communication of positive undertone, found in *deference* and *support*, while low overall pitch is typical for Chinese utterances with a negative undertone, signaling *arrogance*.

9.2.4.4 The role of intonation in Chinese

The role of the Chinese intonation forms defined in Chapter 3 in the signaling of interpersonal attitude is examined in the following order: Mean tonal band-width, final pitch-range, degree of declination and, finally, tonal behavior - first that found on the very last tone or particle of utterance, then the tonal behavior found on longer sequences of utterance-final tones.

(1) The contribution of mean tonal band-width

In the examination of the Chinese suprasegmental profiles, mean tonal band-width was found to vary strongly, not only between different types of interaction strategies but also across the speech samples belonging to one category. These findings rule out the possibility of linking this feature with type of interaction strategy, *valence* or *footing*, a fact which is illustrated in Table 13, showing the mean tonal band-widths of all the Chinese speech samples³⁴.

Searching for another possible correlate of mean tonal band-width, a closer look at column 5 of Table 13 shows that a link between mean tonal band-width and emotional intensity, or arousal, is possible, a phenomenon noted briefly in passing in Chapter 3³⁵. Thus, among Wu's deferent speech samples the utterance with the widest mean tonal band-width, *wLizhi 3*, is the only one signaling a strongly elevated, or high, emotional involvement, while *wNing 4* and *wBing 3*, with narrower mean tonal band-widths, signal slightly elevated emotional arousal, and the last two, *wLizhi 4* and *wLa 4*, with the lowest values for mean tonal band-width are described in terms of low emotional arousal. A similar pattern is found among Wu's *arrogant*

³³ The discussed link between *fluency* and (relative) *footing* is based on less evidence from the Chinese suprasegmental profiles as in the German ones, because many of the Chinese speech samples simultaneously contain different degrees of *hyper-* and *disfluency*, whose overall effect cannot be ascertained.

³⁴ A further, somewhat looser, link is also found in Table 13 between mean tonal band-width and loudness: In a large number of cases, utterances marked by elevated values for loudness also have a wide mean tonal band-width.

³⁵ The emotional intensity signaled by these speech samples, indicated as 'elevated', 'slightly elevated', 'mid' and 'low', was determined in discussion rounds involving some of the native Chinese participants in the phonetic listening tests.

utterances and those judged ambiguous for *valence*: Thus, the 3 *arrogant* speech samples with the highest values for mean tonal band-width, *wBing 2*, (≥ 143 Hz), *wMan 2* (91 Hz) and *wLizhi 2* (196.8 Hz), all signal strongly elevated emotional arousal, while in the ambiguous category, *wLu 4*, holding the lowest value for tonal band-width, is the only one containing mid, i.e. unmarked, emotional intensity, the three other utterances, *wDan 11*, *wMan 3* and *wNing 2*, with wide tonal band-widths (187, >212.3 and 199 Hz, respectively), communicating strongly elevated emotional arousal. Finally, evidence supporting the proposed link is also found in Wu's *supportive* category: Here, the three utterances with the highest values for mean tonal band-width, *wDan 4* (121.6 Hz), *wMan 4* (167.5 Hz) and *wDan 12* (≥ 110.6 Hz) are the only ones communicating a strongly elevated emotional intensity, while the other two, *wBing 1* and *wLizhi 1*, with narrower mean tonal band-widths (76.2 and 106.9 Hz) are described in terms of 'mid' emotional arousal. Coming now to You's speech corpus, the proposed link between mean tonal band-width and emotional intensity also finds support here: Thus, in You's *arrogant* category, the highest value for mean tonal band-width, 136.5 Hz, is found in *yLizhi 2*, communicating a strongly elevated emotional intensity. Slightly narrower mean tonal band-widths, 119.5 and 110.8 Hz, are found in *yDang 1* and *yLong 1*, described in terms of strongly and slightly elevated emotional intensity. The lowest values for mean tonal band-width in this group are in *yMan 2* (≥ 42.5 Hz), *yLong 2* (57.3 Hz.) and *yBing 4* (≥ 60.6 Hz) all three of which signal mid emotional arousal. Similarly, in You's *supportive* group, the highest values for mean tonal band-width (≥ 97.3 , 97, 75.4 and ≥ 78 Hz) are in *yLing 1*, *yLizhi 1*, *yLong 4* and *yMan 1*, all of which signal a strongly or slightly elevated emotional arousal. Finally, You's 4 *deferent* utterances, all of which communicate strongly or slightly elevated emotional intensity, hold values for tonal band-width between 82 and ≥ 122.6 Hz., values comparable with those in her other categories. Thus, while a direct contribution to the signaling of interpersonal attitude must be ruled out in the case of mean tonal band-width, the evidence in the Chinese suprasegmental profiles show that this feature is linked to the emotional intensity signaled by utterances.

(2) *The contribution of final pitch-range*

The examination of the Chinese suprasegmental profiles showed that final pitch-range varied very strongly in the Chinese corpus. No link is apparent between this feature and type of interaction strategy, *valence* or *footing*, and it must be concluded that this feature, at least in the manner defined here³⁶, is not linked in any way to the signaling of speaker attitude.

(3) *The contribution of declination*

According to what appears to be the prevailing view in the literature, declination does not hold communicative functions related to the signaling of attitude. However, in order to examine the phonetic or phonological - and thus potentially communicative - status of this feature, degree of declination was nevertheless determined for as many of the Chinese speech samples as possible. The results obtained in these calculations are summarized briefly below, on the basis of Table 14, where they are displayed together with the syntactic/illocutionary function of each respective utterance - statements, particle questions, non-particle questions and exclamatives³⁷.

The data in Table 14 show that declination varies considerably, not only across the two speakers but also among the utterances produced by one and the same speaker³⁸. A closer look at individual results, however, does suggest the following three broad trends: First, declination in

³⁶ This finding is presumably linked to the omission of final particles in the calculation of final pitch-range.

³⁷ As was discussed in Chapter 3, a correlation between declination and the syntactic function or illocutionary force of utterances has been suggested by some researchers.

³⁸ This variation between individual speakers is the reason why declination is generally examined across the speech of individual speakers.

the statements of both speakers, varying around 10.6 Hz³⁹, is appreciably lower than that in their questions. A mean value of ≥ 21.10 Hz was found for You's non-particle questions, while *wLa 3*, Wu's one particle-question in this table, has a declination of 37.5 Hz. Secondly, the latter statement points to a difference between the two types of questions differentiated here: The data in Table 14 suggest that declination is stronger in questions with sentence particles (in this case *ma*), less strong in questions without sentence particles but containing Chinese *wh*-words, like *wei-shemma* ('why?'), *zemma* ('how?') and *duo-shao* ('how much?'). Finally, Table 13 also suggests that declination is very weak in exclamations, a mean of 2.15 Hz. being calculated for You's two exclamatory utterances. On the basis of these results it may be concluded that while the role of declination in Chinese needs to be specified further, it does appear to be linked to the syntactic and/or illocutionary type of utterances⁴⁰.

(4) *The contribution of final tonal behavior*

In Chapter 3, three types of final tonal behavior, the term used here to describe utterance-final intonation patterns in Chinese, were distinguished: (1) the tonal behavior found on utterance-final sentence particles, (2) the tonal behavior which is sometimes 'added onto' the tones of final (non-particle) syllables, and (3) the tonal behavior traced by the initial pitch, or *range*, of the tones⁴¹ of utterance-final sequences of syllables, beginning on the last strongly stressed syllable (the *focus*) and extending to the last, utterance-final syllable. In Chapter 3, the tonal behavior initiated by the focused syllable was referred to as *focal tonal behavior*, while the tonal behavior found independently of focal tonal behavior on utterance-final tones/particles was referred to as *caudal tonal behavior*.

In the examination of the Chinese speech samples and their suprasegmental profiles it was discovered that types (1) and (2) - the tonal behavior found on utterance-final particles and that found on/added onto the tones of final, non-particle, syllables - may be taken together and treated as one and the same communicative phenomenon - the tonal behavior of utterance-final tones/particles⁴². It was also found that this phenomenon may be equaled with what had been hitherto referred to as *caudal tonal behavior*⁴³. In consequence, the final tonal behavior of Chinese utterances was redefined in terms of (1) *focal tonal behavior* - the tonal behavior initiated by the *focused* syllable, and (2) *caudal tonal behavior* - the various kinds of tonal behavior found on utterance-final tones/particles⁴⁴. While *caudal tonal behavior* is limited to the very last one or two final (particle or non-particle) syllables, *focal tonal behavior* - like the nuclear pitch-patterns of English/German intonation - extends over a sequence of utterance-final syllables. The Chinese suprasegmental profiles were examined for evidence of a link between their final tonal behavior and their attitudinal message - in terms of type of interaction strategy, *footing* or *valence*. In the following, we first briefly note the results obtained for *caudal tonal behavior*, after which we discuss the results obtained in the examination of *focal tonal behavior*.

³⁹ Mean values of 10.1 Hz. (You) and 11.25 Hz. (You) were calculated for statements.

⁴⁰ The above made statements on declination are partly in contradiction to the results of other studies conducted on this phenomenon. In particular, the finding that the declination of the examined Chinese statements is lower than that of the questions goes counter to the research findings generally expounded in the literature. While this discrepancy may also be due to differences in methodology, it must be pointed out that the data discussed here relate first and foremost to *non-particle* questions, which have not received much attention yet, the results obtained for particle questions being too limited to allow generalisation.

⁴¹ As noted in passing in Chapter 3, these amount to variations in tonal band-width.

⁴² In Chapter 3, these were also referred to as segmental and non-segmental intonational pitch-patterns. Strictly speaking, both types are 'added-onto' sentence-final tones - in the first case this happens with, in the second without, sentence particles.

⁴³ This possibility was mentioned briefly at the end of Chapter 3.

⁴⁴ This in effect means that we now distinguish between 3 possible manifestations of caudal tonal behavior: (1) Tonal behavior - or final pitch-patterns - found on/traced by the tones of sentence particles, (2) Tonal behavior 'added-onto' utterance-final, non-particle, tones, and (3) tonal behavior traced by utterance-final, non-particle, tones. Note, however, that the third type is part of focal tonal behavior and as such included in what is referred to here as an 'overall' fall or rise, while the first two types are independent of prior focal tonal behavior and as such comparable to the minor sentence-final rises often found independently of nuclear pitch-patterns in German. See the discussion of *sMango 3* above.

The tonal behavior of final tones/particles (caudal tonal behavior)

In the discussion of the Chinese suprasegmental profiles, the tonal behavior of the final tone/particle was found to vary strongly across the speech samples of one category. The extent of this variation is illustrated in Table 15, where the caudal tonal behavior (of final tones/particles) is shown in column 3. It can be seen that narrow, or *weak*, caudal tonal behavior (of around 1 to 2 *tones*) mid (between 2.5 and 3.5 *tones*) and wide, or *strong*, tonal behavior (of over 4 *tones*) is found in each of the 4 categories of each speaker, so that a link to type of interaction strategy is clearly not possible. Table 15 also shows that a link between caudal tonal behavior and (relative) *footing* does not seem possible either: While Wu's strongest *arrogant* speech sample, *wLizhi 2*, does contain the widest, strongest, caudal tonal behavior of all her utterances (a slope⁴⁵ of 4 *tones* on *yah*), this is not the case in *yLizhi 2*, You's strongest *arrogant* utterance, whose narrow, weak, caudal tonal behavior (a fall of 2.5 *tones* on *qian*), is by far surpassed by that in *yDang 1* (a slope of 5.5 *tones* on *ji*), holding one of the weakest scores for *footing*/strength of position in this category. Similarly strong caudal tonal behavior is also found in *yMing 3*, a *deferent* utterance (a slope of ≤ 5.5 *tones* on *fen*) and in *yLing 1*, a *supportive* utterance (a slope of ≤ 6.5 *tones* on *ke*), the latter also one of the weaker utterances in its category. Likewise, while in Wu's *supportive* group there seems to be negative correlation between relative strength/*footing* and caudal tonal behavior - the stronger utterances *wDan 12* and *wDan 4*, end on weaker caudal tonal behavior (\emptyset and ≥ 0.75 *tones*) than the three weaker ones, *wMan 4*, *wLizhi 1* and *wBing 1* (≥ 1.5 , ≥ 2.5 and ≥ 3 *tones*), this relationship clearly does not apply in her *arrogant* category, where the caudal tonal behavior of the weakest speech sample, *wLa 3* (a fall of ≤ 2 *tones*), is in fact identical to that of the *wBing 2*, with the 2nd highest score for *footing*/strength of position. Finally, while the strongest of You's *deferent* utterances, *yMing 3*, does contain the strongest caudal tonal behavior in this group (a slope of ≤ 5.5 *tones*), the utterance in Wu's *deferent* category with the highest score for *footing*/strength of position, *wLa 4*, ends on the weakest caudal behavior in its group (\emptyset *tones*). On the basis of these data it must be concluded that caudal tonal behavior is not linked in any way to the interaction strategy, *valence* or *footing* signaled by Chinese utterances.

The tonal behavior of utterance-final sequences of syllables (focal tonal behavior)

A somewhat different picture emerges when the tonal behavior of utterances-final sequences of syllables, beginning with *focal tonal behavior*, is taken into consideration. As shown in column 4 of Table 15, in a number of utterances the range, or strength, of focal tonal behavior correlates well with relative strength, or *footing*. Thus, among Wu's 4 *arrogant* utterances, the three strongest, *wLizhi 2*, *wBing 2* and *wMan 2*, all have a wider, stronger, focal tonal behavior - *overall falls* of 12, 7.5 and 9.5 *tones*, followed by a caudal *rise* of 4 and ≤ 2 *tones* in the first two cases - than the weakest in this group, *wLa 3*, which ends on a *rise* of only 5.5 *tones*, followed by a fall of ≤ 2 *tones* on the final particle *ma*⁴⁶. The proposed link also holds among You's *supportive* speech samples, where *yLong 4*, *yLizhi 1* and *yLing 1*, holding the strongest scores for *footing*/strength of position, have strong focal tonal behavior (an *overall fall* of ≤ 5 *tones* followed by an *overall rise* of ≤ 1.5 *tones*; a *fall* of ≥ 4.5 *tones* followed by a *rise* and a *slope* of ≤ 1.5 *tones*; and a fall of 3.5 *tones* followed by a *rise* and a *slope* of ≥ 6.5 *tones*) than *yMan 3* and *yBing 4*, with weaker scores (a fall of 3.5 *tones* followed by a fall of ≤ 1.5 *tones*, and an *overall fall* of ≥ 2 *tones*). The exception to the pattern seems to be *yMan 1*,

⁴⁵ While the term *slope* relates to the pitch-range covered by a given lexical tone, the terms *fall* and *rise* relate to the pitch-range lying between the (initial) pitch of consecutive tones. The terms *overall rise/fall* often used to describe focal tonal behavior (see below) relate to the overall pitch-(range)patterns traced by the initial pitch of consecutive tones.

⁴⁶ Note that it cannot be ascertained at this point whether and in what way caudal tonal behavior which is independent of focal tonal behavior contributes to the attitudinal message of Chinese utterances. At present, it is assumed that focal tonal behavior has a more important, or *primary*, role in these processes, but caudal tonal behavior will be indicated in all cases. A brief discussion of the communicative role of caudal behavior carried by sentence-final particles is included below.

whose strong scores [3.1;+2.4] stand in contrast to its narrow, gentle, focal tonal behavior, a fall of ≤ 2.5 tones, followed by an overall rise of ≤ 3 tones⁴⁷. With the exception of *yDang 1* and *yLong 1*⁴⁸, containing strong focal tonal behavior but weak scores for *footing*/strength of position, the proposed correlation also applies to You's arrogant category. Here, the five strongest speech samples, *yLizhi 2*, *yLa 2*, *yLong 2*, *yMan 2* and *yLu 2*, have stronger focal tonal behavior, i.e. (overall) falls of 5, 3, 3, 3 and 3.5 tones, which in *yLong 2* is followed by a rise of ≤ 1.5 tones, than the weakest utterance, *yLing 2*, with an overall fall of only ≤ 3 tones. Finally, the proposed association between *focal* tonal behavior and (relative) footing also holds among You's *deferent* utterances, where *yMing 3* and *yLizhi 4*, the two strongest, contain stronger focal tonal behavior (a fall of ≥ 5 tones followed by a rise/slope of 5.5 tones, and an overall fall of ≤ 8 Tones), than the weaker two, *yLong 3* (a fall of ≥ 4 tones followed by a rise of 2.5 tones) and *yLu 3* (a rise/slope of ≥ 3.5 tones, followed by a rise/slope of ≥ 1 tone). And among Wu's *deferent* utterances the strongest focal tonal behavior (an overall fall of ≤ 4 tones and a fall of 4.5 tones followed by a fall of 1.5 tones) is found in *wLizhi 3* and *wBing 3*, the strongest utterances here, while the weakest speech samples in this category, *wNing 4* and *wLizhi 4*, end on a weaker focal tonal behavior - a fall of ≥ 3.5 tones and a rise of ≤ 1 tone and a fall of ≥ 2 tones followed by a rise of 2 tones and a slope of 1 tone.

Thus, as a first result in the current examination, the Chinese suprasegmental profiles indicate a link between the (relative) *footing* signaled by the Chinese speech samples and the strength, in terms of the range, of the focal tonal behavior they contain.

In order to gain a clearer understanding of the contribution of tonal behavior as compared to that of the other suprasegmental features under discussion here, we now take a closer look at the suprasegmental profiles of those speech samples in which the proposed link was not found to hold. First of all, it was noted above as rather puzzling that of You's strongly *arrogant* speech samples *yLong 2* received a slightly higher score for *footing* than *yMan 2*, despite the fact that the focal tonal behavior of both utterances is largely identical, involving falls of about 3 tones. A look at their suprasegmental profiles shows that *yLong 2* has elevated values for loudness and tempo together with *fluency*, and *yMan 2* has mid to elevated loudness/tempo and also *fluency*, so that the communicative effect of their overall features is also largely the same. On the other hand, however, the phonatory *settings* of these two speech samples are quite different: While *yLong 2* contains the exclusive presence of *modal-tenseness* and *tenseness*, *yMan 2* is largely *modal*, alternating only once with *modal-tenseness*. This suggests that the difference in scores between these two utterances was caused not by their final tonal behavior but rather by this difference in phonatory voice quality.

As the second case to be examined more closely, among Wu's 3 strongly *arrogant* utterances - *wLizhi 2*, *wBing 2* and *wMan 2* - it is *wBing 2* which holds the 2nd strongest overall scores, despite the fact that its focal tonal behavior (an overall fall of 7.5 tones followed by a caudal rise of ≤ 2 tones) is not stronger by weaker than that, for example, of *wMan 2* (an overall fall of 9.5 tones), which received the 3rd strongest score. A look at the suprasegmental profiles of these two utterances shows that *wBing 2* is marked by elevated values for loudness, tempo and fluency and the almost exclusive presence of *tenseness*, while *wMan 2*, which also has similarly elevated values for the overall features, contains the almost exclusive presence of (modal-)laxness. This suggests that the marked presence of *tenseness* in *wBing 2* contributed more strongly to the high scores this utterance received for *footing*/ strength of position than its focal tonal behavior. On the other hand, a look at the caudal tonal behavior of *wBing 2* shows that it is a (low) rise of ≤ 2 tones (on the final particle *ah*) which strongly resembles the low rise of English/German intonation, associated above with a strong, potentially negative

⁴⁷ This phenomenon will be discussed in more detail below.

⁴⁸ More will be said of these two speech samples, *yLong 1* and *yDang 1* below.

speaker attitude. It is not inconceivable that this (*caudal*) *low rise* on *ah* in *wBing 2* has a similar communicative effect in Chinese and that it thus also contributed to the strong score this utterance received in the judgement tests.

Further evidence for the strong, potentially negative effect of such a (*caudal*) *low rise* is found in two of Wu's *arrogant* speech samples, *wLizhi 2* and *wBing 2*, where it is found on the particles *ah* and *yah*. The thing that is particularly striking about these two utterances is that they hold much stronger scores for *footing*/strength of position than *wMan 2* and *wLa 3*, the other two speech samples in this category - [4.8;-1.8] and [4.5;-2.1], as compared to [3.8; -1.9] and [3.3; -1.9]. Even if one allows for the fact that *wLizhi 2* and *wBing 2* contain a stronger presence of *tenseness*, a slightly stronger combination of overall features and wide (*overall*) *focal falls* of ≤ 11 and 7.5 *tones*, respectively, the difference in the scores of these four utterances is still rather striking, as *wMan 2* and *wLa 3* also have rather wide (*overall*) *focal falls* of 9.5 and ≥ 5.5 *tones*. A closer look at the caudal tonal behavior of these speech samples shows *wLizhi 2* and *wBing 2* are special in that they contain *caudal low rises* - of 4 *tones* in *wLizhi 2* and ≤ 2 *tones* in *wBing 2* - which are carried by the final particles *ah* (in *wBing 2*) and *yah* (in *wLizhi 2*). No such *caudal low rises* are found in *wMan 2* and *wLa 3*.

Finally, it was also observed above that *yMan 1* [3.1; +2.4] is among the strongest of You's *supportive* utterances, despite the fact that its focal tonal behavior, a *fall* of ≤ 2.5 *tones*, followed by a *rise* of ≤ 3 *tones*, is relatively narrow/weak. In fact, it particularly is striking that this utterance, containing much the same phonatory *settings* as *yLong 4* but a weaker combination of overall features, actually received the same overall scores as *yLong 4* [3.3; +2.2]. A closer look at the caudal tonal behavior of these two utterances shows that *yLong 4* contains a *caudal (overall) low rise* of only ≤ 1.5 *tones* (carried by *pian*^{2°}, *yi*^{3°} and *oh*), while *yMan 1* has a *caudal (overall) low rise* of ≤ 3 *tones* (carried by *ming*², *de*[°] and *ah*). This suggests that not only the direction but also the range covered by *caudal low rises* contributes to the relative *footing*/strength of position signaled by Chinese utterances⁴⁹.

As a final point to be added here, in the light of the above findings it now seems very likely that the different scores of *wMan 2* [3.8;-1.9] and *wLu 4*, [2.6; +/- 2.0], to which attention was briefly called in Chapter 3, were indeed caused by their *caudal* tonal behavior (on *ba*). A closer look at their suprasegmental profiles shows that while *wMan 2* has a slightly stronger presence of *modal*- and *modal-tenseness* and a stronger combination of overall features than *wLu 4*, a fact which may partly account for its higher score, in fact both utterances have similarly strong focal tonal behavior - a *fall* of 9 *tones* in *wLu 4* and a *fall* of 8.5 *tones* in *wMan 2*. A look at their caudal tonal behavior, however, shows that in *wMan 2* this is a *fall* of ≤ 2 *tones* which concludes the deep focal fall of this utterance, while in *wLu 4* it is a narrow, gradual fall of 1.5 *tones*, whose auditory effect may be described as gentle, and which may conceivably have assuaged the strong, forceful, effect of the preceding *focal tonal behavior*. Interestingly, this high narrow caudal fall seems very similar to the high, narrow, *leveling-off fall* of German intonation, whose communicative effect it also seems to share.

To conclude the above discussion, the evidence found in the Chinese suprasegmental profiles indicates that focal tonal behavior - i.e. utterance-final intonation - in Chinese makes a strong contribution to the signaling of speaker attitude. While the effect of other suprasegmental features, such as phonatory voice quality and the overall features must also be considered, there is clearly a strong link between the range (or strength) of the focal tonal behavior and the (relative) *footing* communicated by Chinese utterances: Strong focal tonal behavior, covering a wide range, is found in utterances signaling a strong, superior *footing*, while weak, narrow focal tonal behavior is a feature of utterances signaling a weaker, inferior speaker

⁴⁹ As an additional point of interest, this caudal tonal behavior is not only traced by the tones of final particles, such as *ah*, *oh* and *yah*, but also by other utterance-final tones,

attitude. As a second result, while the role of focal tonal behavior in these signaling processes is clearly more important than that of caudal tonal behavior, the suprasegmental profiles do suggest that certain types of caudal tonal behavior also contribute strongly to the signaling of (relative) *footing*. As two examples, the *caudal low rise* - carried either by a final particle or a sequence of final tones - was found to be associated with a strong, potentially negative attitude, precisely the communicative effect which is associated with the *low rise* of the English/German intonation system. Evidence was also found that a high, narrow and gently descending *caudal fall* communicates a similarly gentle interpersonal message as the high, narrow leveling-off fall of German intonation.

9.2.4.5 The role of supralaryngeal voice quality

In the Chinese speech samples two supralaryngeal *settings*, *palatalization* and *labiodentalized voice* were detected. As *delayed glottalization* sometimes accompanies *palatalization*, the discussion of this phenomenon, strictly speaking a feature of phonatory vibration, will also be included in this subsection. As shown in Table D (Section 2) of Appendix III, *palatalization* together with *delayed glottalization* is found in *wMing 4* and *yLong 1*. *Delayed glottalization* on its own, on the other hand, was detected in two utterances, *yMing 3* and *yLing 1*. In order to determine the communicative effect of these *settings* the suprasegmental profiles of these speech samples follows, taking into account the communicative effect of the other suprasegmental features, as determined above.

The first of these speech samples, *wNing 4* [1.8; +1.8] is the weakest among Wu's *deferent* utterances. It contains mid to reduced values for loudness, tempo and fluency and the almost exclusive presence of *laxness*, developing into *breathiness*. On the other hand, it has the 2nd strongest focal tonal behavior in its category, a *fall* of ≥ 3.5 tones (followed by a *caudal rise* of ≤ 1 tone). Taken together, this adds up to two weak features (phonatory voice quality and the overall features) but also one rather strong feature, the final tonal behavior. As this utterance holds the weakest score for *footing*/strength of position not only in Wu's *deferent* category but in her entire speech corpus, this suggests that - similar to what was found for the German speech samples - the effect of *palatalization* (and/or *delayed glottalization*) was weak and/or positive here, supporting the weak and positive effect of the phonatory *settings* and the overall features. Coming now to *yLong 1* [3.6; -2.1], this is one of the weaker utterances in You's *arrogant* group. It contains *modalness* alternating with *tenseness*, mid tempo/amplitude and *fluency* and a very strong final tonal behavior, a *focal fall* of 5.5 tones (followed by a rise of 4 and a slope of ≤ 4.5 tones). While the phonatory *settings* and the overall features are likely to have contributed to the strong scores for strength of position and/or the negative *valence* of this speech sample, on the other hand this speech sample have relatively weak scores for *footing*, despite its strong final tonal behavior. Therefore it seems very likely that here too, the effect of *palatalization* (and/or *delayed glottalization*) was weak/positive, contributing to its weak score for *footing*. However this explanation is not completely satisfactory, since the scores of this utterance are remarkably weak given the marked presence of *tenseness* here and the absence of other suprasegmental features, apart from *palatalization*, associable with the signaling of weak *footing*/strength of position. More is said of this speech sample below.

Coming now to *yMing 3* and *yLing 1*, *delayed glottalization* was found alone here, without *palatalization*. The first of these two utterances, *yMing 3*, is the strongest of You's *deferent* speech samples, with a score of [3.1; +1.5]. It is marked by a very strong final tonal behavior (a *focal fall* of 5 tones (followed by a rise/fall of ≤ 5.5 tones), mid to reduced loudness, tempo and fluency and a strong presence of *modalness* and *modal-tenseness*. Given the presence of two strong elements (phonatory voice quality and final tonal behavior) but only one weak feature (the overall features), it seems highly possible that *delayed glottalization* had a weak, perhaps also positive, communicative effect here. Indeed, given the virtual absence of warm,

lax phonatory *settings* in this utterance, it is very possible that delayed *glottalization* not only lowered the scores for footing/strength of position but also had the effect of adding a somewhat more gentle undertone, without which *yMing 3* may not have ‘qualified for’ the *deferent* category. We will return to this utterance below. Finally, *yLing 1* [3.2;+2.1] is one of You’s weaker *supportive* speech samples. It has a relatively strong final tonal behavior, a *focal fall* of ≥ 3.5 tones (followed by a rise/slope of ≤ 6.5 tones), mid tempo, amplitude and fluency and the strong presence of *modal-* and *modal-tenses* - all features which were established above to contribute to the signaling of relative strength and, in the case of the phonatory voice qualities, also negativeness. Given this preponderance of strong, potentially negative suprasegmental features, the relatively weak and positive scores of this utterance suggest that, here too, the effect of delayed *glottalization* was to lower the scores for *footing* and add a gentler, more positive undertone to the utterance.

Turning now to *labiodentalized voice (smiling)*, auditory evidence for this *setting* was found in *yLizhi 4* and *yBing 4*. *YLizhi 4*, with a score of [2.7; +1.9], is one of the stronger utterances in You’s *deferent* category. It is marked by the almost exclusive presence of *tenseness*, 2nd degree *disfluency*, mid tempo/loudness and has the strongest tonal behavior here, an overall focal fall of ≤ 8 tones. As this utterance contains two very strong/ negative features (*tenseness* and the tonal behavior) and only one weak/positive feature (the overall features), its positive *valence* is somewhat puzzling. It seems very plausible that the presence of *labiodentalization (smiling)* added an additional positive effect, allowing this utterance to be judged in terms of *deference*. Similarly, *yBing 4* [2.9; +1.9] is the weakest of You’s *supportive* utterances. It has mid tempo, loudness and *fluency*, a relatively weak tonal behavior (a *focal overall fall* of ≥ 2 tones) and *tense* voice quality, developing into *laxness* towards the end. While the relatively strong presence of *tenseness* and the lack of expressedly weak overall features may be linked to the relatively high scores for strength of position, and the relatively low score for *footing* may be associated with the weak final tonal behavior, once again the positive *valence* of this utterance is rather puzzling, unless one assumes that the presence of *labiodentalized voice* had this positive communicative effect.

To conclude, the evidence in the Chinese speech samples indicates that *palatalization* and delayed *glottalization* may be associated with the signaling of relative weakness, or inferiority, while *labiodentalization* contributes to the positive undertone/*valence* of utterances.

9.2.5 Summary: The phonetic exponency of *deference*, *arrogance* and *support* in Chinese

Based on the above results, the phonetic exponency of *deference*, *arrogance* and *support* in Chinese can now be defined more precisely.

Deference

The above results indicate that *deference* in Chinese is communicated by (*modal-*) *lax*, *lax* and/or *breathy* voice quality. *Palatalization* and delayed *glottalization* are also frequently encountered here, but *modal voice* and other *tenser* phonatory *settings* are only found in utterances communicating a strong *footing*. *Deference* is further signaled by mid to reduced, only rarely elevated, loudness, tempo and fluency. Reduced values for more than one of these phenomena signal a weak, inferior *footing*, while elevated values are characteristic of utterances communicating a stronger, more forceful attitude. With regards to intonation⁵⁰, the weak, inferior *footing* typical of *deference* is communicated by weak focal tonal behavior, covering a narrow range. A stronger, more forceful attitude is signaled by focal tonal behavior with a

⁵⁰ As these two phenomena are not directly linked to the signaling of attitudinal messages, the results for declination and mean tonal bandwidth will not receive further mention here.

wider range. Finally, mid to high overall pitch and *labiodentali-zation* (smiling), contributing to the signaling of positive *valence*, are also encountered in the communication of *deference*.

Arrogance

The results obtained above indicate that *arrogance* in Chinese is communicated by *modalness* and *modal-tenseness*. *Tenseness* on its own is found in strongly arrogant utterances, while *laxer*, softer phonatory *settings*, such as (*modal-*) *laxness*, are found in less strongly arrogant utterances. On the other hand, *palatalization*, delayed *glottalization* and *labiodentalization* are usually not encountered in this interaction strategy. *Arrogance* is also signaled with mid to elevated loudness, tempo and fluency. Elevated values for more than one of these phenomena are found in strong arrogance, while mid, and occasionally also reduced, values for the overall features are common in utterances signaling a weaker, less forceful *footing*, or attitude. With regards to intonation, the strong, superior attitude typical of arrogant utterances is signaled by wide focal tonal behavior, covering a wide range, while weak, i.e. narrow, focal tonal behavior is a feature of utterances signaling a weaker, less forceful speaker attitude. Finally, the negative undertone of *arrogance* may also be communicated by low overall pitch.

Support

The above results indicate that *support* in Chinese is communicated by *modal-laxness*, *lax-* and/or *breathiness*. Utterances often begin with *tenser* phonatory *settings*, which then develop into *laxer*, softer, *settings* by mid-utterance. *Support* is further signaled by mid tempo, loudness and fluency, elevated values for one or more of these phenomena found in utterances signaling a strong *footing*, while reduced values are characteristic of utterances signaling a weaker, inferior, attitude. With regards to intonation, focal tonal behavior is wide and strong in utterances signaling a strong, forceful *footing*, less wide in utterances communicating a gentler, less forceful attitude on the part of the speaker. Given its link to positive *valence*, overall pitch is usually relatively high in supportive utterances. Finally, *labiodentalization* and delayed *glottalization*, the latter often found together with *palatalization*, are also found here.

9.2.6 The Chinese utterances signaling ambiguous *valence*

Five Chinese speech samples were judged to be ambiguous for *valence*: *yMing 1*, *wNing 2*, *wMan 3*, *wLu 4*, *wDan 11*. The results of their comparative examination are as follows:

As a first point, it was found that the suprasegmental profiles of the Chinese speech samples with ambiguous *valence* support the statements made above on the phonetic exponency of *deference*, *support* and *arrogance* in Chinese and the role of the suprasegmental features in these signaling processes. Strong support in particular was found for the above defined role of the overall features and the phonatory *settings*, not only in the communication of *valence/undertone* but also in the signaling of (relative) *footing*. As a second result, the combinations of phonatory *settings* and the overall features in all these speech samples conformed to the patterns defined above, a discovery which contrasts with what was found for the German ambiguous utterances. On the other hand, all but one of these utterances is marked by a strong contrast between their relatively *lax*, i.e. soft and gentle, phonatory *settings* and their very wide, strong tonal behavior, a focal tonal behavior of 8 to 10 *tones* in Wu's speech samples (followed in *wNing 2* by a wide falling added-on pitch-pattern) and a wide caudal low rise of 5 *tones* in You's utterance. It is very likely that this contrast between the gentle message of the phonatory *settings* and the overall features and the strong, forceful message of intonation contributed to the ambiguous judgements of these utterances. On the other hand, given the fact that these contrasts led to ambiguous scores not of *footing* but of *valence*, this finding also indicates that focal tonal behavior can also contribute to the signaling of undertone in Chinese. As a third result, it was found that none of the Chinese utterances in this category contains

palatalization or *labiodentalized voice*, a fact which points to a strong, in the sense of clear and unequivocal, communicative effect on the part of these *supralaryngeal settings*, a finding also noted above for the German speech samples.

As a fourth result of a slightly different nature, two of these ambiguous speech samples, *yMing 3* and *wNing 2*, were discovered to signal two particular interaction strategies hitherto not included in the study. In *yMing 3*, this was the teasing humour strategy of *tai-gang*, signaled here by the pronounced presence of *tense* phonatory *settings*, mid to elevated values for overall loudness, tempo and fluency, a very wide mean tonal band-width and a strong final tonal behavior. A comparative examination of *yDang 1* and *yLong 1*, containing the same combinations of suprasegmental features which were noted above to be rather incompatible with their scores - in particular for *valence*, revealed that these two speech samples were also judged in terms of *tai-gang* in the judgement tests, a finding which supports the association of these features with this interaction strategy⁵¹.

Coming now to *wMing 2*, it will be recalled that in the judgement tests the speech behavior of Wu here was described by a number of the judges as petulant and childish⁵². As such her behavior is highly reminiscent of what is commonly referred to as *sa-jiao*, a wailing, whimpering exaggerated manner of speaking with which the speaker, usually a child or a woman, tries to get his or her way with the listener⁵³. Interestingly, the suprasegmental features in this utterance are almost identical to those found in *tai-gang*, the only difference being the added-on falling pitch-pattern, to which attention was called in passing above. This suggests that this falling added-on pitch-pattern - together with the marked presence of (*modal-*) *tenseness*, the mid to elevated values for loudness, tempo and fluency, the wide mean tonal band-width and the equally wide, and strong, final tonal behavior - serves to mark this interaction strategy⁵⁴.

Finally, a point to which attention has been called before, in contrast to the ambiguous German utterances, none of the Chinese speech samples in this category was found to be associable with *false deference*. This may be regarded as evidence for the strong conventionalization of *deference* in Chinese society, as a result of which utterances whose suprasegmental profiles do not strictly conform to the phonetic exponency of *deference* are nevertheless understood in terms of this strategy.

9.3 Comparison: The signaling of speaker attitude in German and Chinese

In this last section of the present chapter, the manner in which *deference*, *arrogance* and *support*⁵⁵ are signaled in German and Chinese - their phonetic exponency - is discussed comparatively and differences between the two languages are pointed out. Following this, we will consider the manner in which the suprasegmental phenomena contribute to the signaling of speaker attitude in German and Chinese.

9.3.1 The signaling of deference, support arrogance and false deference in German and Chinese

Deference

In both languages *deference* is communicated by (*modal-*)*lax* and/or *breathy* phonatory voice quality, *tenser* phonatory *settings* found in utterances signaling a strong *footing*. The use of the

⁵¹ In the sociolinguistic judgement tests a large number of the judges, especially those who were familiar with You, judged these 3 speech samples as representative of joking irony and/or *tai-gang*: Of the 14 judges who participated in You's judgement test, 10 saw *yDang 1* as representing *tai-gang*, 8 judged *yLong 1* in terms of this strategy and 5 judged *yMing 1* in terms of joking irony/*tai-gang*.

⁵² In fact, 12 out of Wu's 14 judges judged this utterance to represent childish speech behavior, with a weak/submissive *footing* and either positive or ambiguous *valence*, a negative undertone indicated in only two cases.

⁵³ The translation given for *sa-jiao* in the New Practical Chinese English Dictionary is: 'to pretend to be angry or displeased, of a woman'.

⁵⁴ Chao's (1968) meanings for added-on *falls* includes the display of 'pretended - in the sense of exaggerated - emotionality', a connotation which fits well with the judges' description of Wu's speech behavior here.

⁵⁵ As no speech samples were unequivocally allocated to *false deference* this strategy will not receive further consideration here.

supralaryngeal *setting* of *palatalization*, optionally accompanied in Chinese by delayed *glottalization*, was also observed in both languages. *Deference* is further signaled by mid to reduced, only rarely elevated, loudness, tempo and fluency. Reduced values for more than one of these phenomena signal a weak, inferior *footing*, while elevated values are common in utterances communicating a stronger, more forceful attitude. Mid to high overall pitch is also typical of deferent utterances in both languages. This feature is linked to the signaling of weak *footing* in German, to the communication of positive *valence* in Chinese. With regards to intonation, the weak, inferior *footing* typical of *deference* is communicated in German by narrow, and often high, nuclear pitch-patterns, in Chinese by weak, narrow, focal tonal behavior. A stronger, more forceful attitude is signaled by wider nuclear pitch-patterns and wider focal tonal behavior. A further feature of *deferent* utterances in German is their relatively narrow prenuclear pitch-range, linked to the positive *valence*/undertone of this strategy.

Support

In Chinese and German *support* is communicated by *modal-laxness*, *lax-* and/or *breathiness*, the latter more frequent in Chinese than in German, where *breathiness* appears to be more typical of *deference*. In Chinese, utterances often begin with a relatively *tense* phonatory *setting*, which then develops into *laxer*, softer, often *breathy* voice quality by mid-utterance. Interestingly, *palatalization*, frequent among the deferent utterances, was not found in any supportive speech samples in the corpus, a finding which suggests that it can be used to distinguish between *deference* and *support* in both languages. *Support* is further signaled by mid tempo, loudness and fluency, elevated values for one or more of these phenomena found in utterances signaling a strong *footing*, while reduced values are characteristic of utterances signaling a weaker, inferior, attitude on the part of the speaker. Conversely, a weak *footing* is signaled in German by a high, elevated, overall pitch, a lower pitch being more common in the communication of a strong, forceful attitude. Given its link to positive *valence* in Chinese, overall pitch is generally relatively high in Chinese supportive utterances. With regards to intonation, nuclear pitch-patterns and focal tonal behavior are wide and strong in supportive utterances signaling a strong, superior *footing*, less wide in utterances communicating a gentler, less forceful attitude. In German, an additional feature of *support* is relatively narrow prenuclear pitch-range linked to the positive *valence*/undertone of this strategy.

Arrogance

In both languages *arrogance* is signaled by the pronounced presence of *modalness*, blending frequently with *modal-tenseness* and/or *tenseness*, only rarely with *laxness* or *modal-laxness*. *Tenseness* on its own is found in strongly arrogant utterances, while *laxer*, softer phonatory *settings* are common in less strongly arrogant utterances. *Palatalization* is not employed in this interaction strategy. In both languages *arrogance* is further signaled by mid to elevated loudness, tempo and fluency. Elevated values for more than one of these phenomena communicate strong *arrogance*, while mid, and occasionally also reduced, values for the overall features are common in utterances signaling a weaker, less forceful *footing*, or attitude. In both languages, an additional feature of *arrogance* is a relatively low overall pitch, linked in German to the signaling of a strong, forceful attitude, in Chinese to the negative *valence*/undertone characteristic of this strategy. With regards to intonation, the strong, superior *footing* typical of arrogant utterances is signaled in German by wide nuclear pitch-patterns and in Chinese by wide focal tonal behavior, weaker, i.e. narrower, nuclear pitch-patterns and focal tonal behavior used to communicate a weaker, less forceful speaker attitude. Finally, the negative undertone of *arrogance* is also communicated in German by a relatively wide prenuclear pitch-range.

9.3.2 The role of the suprasegmental features in German and Chinese

9.3.2.1 The role of phonatory voice quality in German and Chinese

In German and Chinese phonatory voice quality was found to be associated primarily with the *valence*, or undertone, of utterances. Broadly speaking, *modal-tenseness*, alternating with *tenseness* or *modalness* communicates negative *valence*, while positive *valence* is signaled by the presence of softer phonatory *settings*, i.e. (*modal*-) *laxness* and *breathiness*. As a second function, in both languages the phonatory *settings* were also found to contribute to the signaling of relative strength/*footing*: Utterances containing a more pronounced, presence of *tenseness* and *modal-tenseness* are often among the stronger utterances in one category, speech samples containing a stronger presence of *laxness* and *breathiness* typically among the weaker utterances in one category⁵⁶. Finally, it was found that in German, but not in Chinese, *breathiness* may be used in the differentiation between *deference* and *support*, since this *setting* is found in German deferent utterances but not in supportive ones. This latter finding also suggests that Chinese women employ the *setting* of *breathiness* more often in social intercourse than German women, where *breathiness* is limited to deferent utterances.

9.3.2.2 The role of overall tempo, loudness and fluency in German and Chinese

In both German and Chinese, (combinations of) the overall features loudness, tempo and fluency were found to be linked to the signaling of (relative) *footing*. Broadly speaking, among the speech samples within one category, i.e. signaling one type of interaction strategy, those containing elevated values for one or more of these phenomena signal a strong, forceful *footing*, while those containing reduced values communicate a less forceful, weaker attitude on the part of the speaker. This association with the relative strength of utterances also leads an indirect link between these features and interaction strategy: As arrogance is generally marked by a strong *footing*/superior attitude, elevated values for loudness, tempo and fluency are more typical of arrogant utterances. Conversely, due to the characteristically weak, inferior *footing* of *deference*, reduced values for these features are generally more common in *deferent* utterances. As a further result, the evidence obtained for *hyper**- and *disfluency**, caused by marked types of utterance-final slow-down, suggests that both these features contribute to the signaling of weak, inferior *footing*, so that they may be expected to be more common in deferent and supportive utterances than in arrogant ones.

9.3.2.3 The role of overall pitch in German and Chinese

The analysis of the German suprasegmental profiles revealed strong evidence for a link in this language overall pitch and (relative) *footing*: German utterances signaling a weak, inferior *footing* are generally marked by a higher overall pitch than utterances signaling a stronger, more forceful attitude on the part of the speaker. Interestingly, a somewhat different link was found in the Chinese corpus, where the relatively high overall pitch of deferent and supportive utterances is linked to the positive *valence* of these strategies, while the negative *valence* of *arrogance* is signaled by low overall pitch.

9.3.2.4 The role of supralaryngeal voice quality in German and Chinese

In both German and Chinese, the suprasegmental profiles point to a link between *palatalization* and relative weakness, or inferior *footing*⁵⁷. The examination of the Chinese profiles showed that *labiodentalized voice* not only contributes to the signaling of relative weakness in this language, but potentially also to the communication of positive *valence*. Interestingly, none of these *settings* were present in the speech samples with ambiguous *valence*, suggesting

⁵⁶As a further point mentioned in passing above, *modal voice* in German was found to contribute more to the communication of a relatively strong, potentially negative, attitudinal message than attitudinal neutrality.

⁵⁷In the Chinese speech samples, *palatalization* is often accompanied by delayed *glottalization*, indicating a similar function for this feature.

that their communicative effect is relatively clear and unambiguous. Finally, the fact that *deference* is frequent among the deferent utterances of the 4 speakers indicates that this *setting* may also be regarded as a characteristic feature of this interaction strategy in both languages.

9.3.2.5 *The role of intonation in German*

The examination of the German suprasegmental profiles showed that a number of German intonational features may be associated with the signaling of relative strength or *footing*. Thus, wide and, especially, *steep*, falls and rises communicate a relatively strong, forceful and potentially superior, *footing*, while mid and, especially, narrow nuclear pitch-patterns signal a gentler, more inferior attitude on the part of the speaker, particularly when their *tail slopes* are gradual and gentle. The weakest German nuclear pitch-patterns were found to be high narrow falls, described as gradually *leveling off* to high or mid-high pitch. On the other hand, a strong link was also found between wide and steeply, or abruptly, descending falls and a strong, forceful attitude. The strongest, most negative attitudinal messages were found to be signaled by wide and abruptly falling nuclear pitch-patterns, such as those containing mid or low ‘drop’ nuclei, and by *quasi-stylization*, in which the mid/low ‘drop’ nucleus is followed by a *flat tail* and preceded by a high/mid level *head*, optionally preceded by a low *prehead*. Evidence was also found in support of an association between the low, wide rise and a strong and potentially negative attitudinal message, as proposed in the literature. Coming now to the prenuclear pitch patterns, the frequent presence of low to mid *heads* in the *deferent* and *supportive* German speech samples and mid to high *heads* in the arrogant speech samples indicates a link between wide prenuclear range and negative *valence* and between narrow prenuclear pitch-range and positive *valence*. Finally, repeated evidence was found in support of an association between *glissando heads* and a high degree of strength and potentially also negative *valence* in German, as has been reported for English.

9.3.2.6 *The role of intonation in Chinese*

Turning now to the contributing of Chinese intonation, the evidence in the Chinese suprasegmental profiles indicates that focal tonal behavior - i.e. utterance-final intonation - in Chinese has an important role in the signaling of speaker attitude. While the effect of other suprasegmental features, such as phonatory voice quality and the overall features must also be considered, a strong link was found between the range (or strength) of the focal tonal behavior and the (relative) *footing* communicated by Chinese utterances. Thus, strong focal tonal behavior, covering a wide range, is found in utterances signaling a strong, superior *footing*, while weak, narrow focal tonal behavior is a feature of utterances signaling a weaker, inferior speaker attitude. As a second result, while the role of focal tonal behavior in these signaling processes is clearly more important, more *primary*, than that of caudal tonal behavior, the suprasegmental profiles suggest that certain types of caudal tonal behavior also contribute strongly to the signaling of (relative) *footing*. As two examples, the *caudal low rise* - carried either by a final particle or a sequence of final tones - was found to be linked to a strong, potentially negative attitude, precisely the communicative effect which is associated with the *low rise* of the English/German intonation system. Evidence was also found that high, narrow and gently descending *caudal falls* communicate a similarly gentle interpersonal message as the high, narrow *leveling-off* falls of German intonation.

9.3.3 *Final summary and conclusions*

The findings of the above discussion may be summarized as follows. As a first point, the phonatory *settings* and the overall features clearly hold virtually the same attitudinal functions in German and Chinese, contributing in almost identical manner to the signaling of *valence/undertone* and the (relative) *footing* of utterances. The only marked differences between the

two languages were found with respect to the use of *breathiness* and overall pitch. Whereas *breathiness* is a feature of both *support* and *deference* in Chinese, it is clearly less common in German, where it is limited to the strategy of *deference*. And while overall pitch is linked to *footing* in German, it is associated with *valence*/undertone in Chinese.

As a second point, the supralaryngeal *settings* of *palatalization* and *labiodentalization* were also found to hold very similar functions in German and Chinese - the signaling of relative weakness and in the latter case potentially also positive *valence*/undertone. As *palatalization* is found most frequently in the deferent utterances of the speech corpus, it may also be regarded as a marker of the *deferent* strategy, a statement which also applies to both languages.

Finally, turning to the contribution of intonation, though the examinations did not include Chinese counterparts of German and English *heads*, the results of the above analyses showed that the nuclear pitch-patterns of German (and English) intonation and their Chinese counterparts, the focal tonal behavior of Mandarin Chinese utterances, hold the same attitudinal functions - the signaling of the relative strength, or *footing*. Given the fact that these forms of intonation are signaled, or created, in a rather different manner - involving variations in pitch in German, but variations in pitch-range in Chinese - this is surprising. Evidence for a functional similarity was also found between high, narrow leveling-off pitch-patterns in German and similarly narrow leveling-off caudal tonal behavior in Chinese, both of which signal a weak, inferior *footing*, and between German wide low rises and Chinese low rising caudal behavior, both signaling a strong, potentially negative attitudinal message.

These results indicate that *valence* and *footing*, the components of *deference*, *support* and *arrogance*, are signaled in widely the same manner in German and Chinese, the only differences found between the two languages relating to the use of *breathiness* and overall pitch.

Table 12 Final results for overall loudness, overall tempo and fluency

Clara

Deference	Speech sample	score	Overall loudness	Overall tempo	Fluency
	cBlum 4	1.8; +1.0	soft	mid-slow	disfluent* (1/2)
	cGans 1	2.6; +1.5	mid	slow	disfluent (1)
	cKuli 1	2.3; +1.3	mid	mid-slow	disfluent (1)
	cZwieb 3	1.7; +1.1	mid	mid	disfluent (4)
	cMoon 1	2.5; +1.6	mid	mid	disfluent(1); hyperfluent* (1/2)
	cGans 4	2.9; 1.6	mid	mid-slow	disfluent (2)
Arrogance	Speech sample	score	Overall loudness	Overall tempo	Fluency
	cKuli 3	3.8; -2.2	mid-loud	mid-fast	fluent
	cMango 1	4.1; -1.9	mid-loud	mid-fast	hyperfluent (1)
	cMango 2	3.1; -1.7	mid-soft	mid	disfluent (1)
	cBlut 4	3.5; -1.7	mid	mid	fluent
Support	Speech sample	score	Overall loudness	Overall tempo	Fluency
	cMoon 3	3.2; +2.0	mid	mid-slow	disfluent (1)
	cBlum 3	3.3; +1.9	mid	mid-slow	disfluent (1)
	cMango 3	3.1; +1.5	mid-soft	mid	fluent
	cSchoko 3	3.7; +2.3	mid	mid	fluent
	cBlum 1	3.8; +1.9	mid	slow	fluent
Ambiguous	Speech sample	score	Overall loudness	Overall tempo	Fluency
	cSchoko 2	2.3; +/-1.1	mid-soft	mid-fast	fluent
	cKuli 2	2.8; +/-1.5	mid-soft	mid-fast	disfluent(1); hyperfluent* (1/2)
	cBlut 3	3.5; +/-2.0	mid	fast-slow	fluent
	cVani 2	3.9; +/-2.1	mid-soft	mid-fast	fluent
	cZwieb 2	4.0; +/-2.2	mid	fast-mid	fluent

Sybillie

Deference	Speech sample	score	Overall loudness	Overall tempo	Fluency
	sKuli 1	2.4; +1.2	mid-loud	mid	hyperfluent* (1/2)
Arrogance	Speech sample	score	Overall loudness	Overall tempo	Fluency
	sKiwi 1	3.5; -2.0	mid	mid	hyperfluent (1)
	sKiwi 2	3.6; -1.7	mid	fast	hyperfluent (1)
	sMango 3	4.4; -2.3	mid	fast	hyperfluent (1)
	sSchoko 1	3.4; -2.1	mid	mid	disfluent (2)
Support	Speech sample	score	Overall loudness	Overall tempo	Fluency
	sMango 4	2.8; +1.9	mid-loud	mid	hyperfluent (1)
	sKiwi 4	3.1; +2.0	mid	fast	fluent
	sMoon 1	3.4; +1.9	mid	mid-fast	hyperfluent* (1/2); hyperfluent (1)
	sBlum 3	3.5; +2.0	mid	fast	fluent
	sKuli 4	3.8; +2.2	mid	fast	hyperfluent (1)
	sKiwi 3	3.6; +2.3	mid	mid	fluent
Ambiguous	Speech sample	score	Overall loudness	Overall tempo	Fluency
	sGans 1	2.9; +/-1.7	mid	mid-fast	fluent
	sBlum 1	3.1; +/-1.2	mid	mid-fast	hyperfluent* (1/2)
	sMoon 2	3.6; +/-2.1	mid	mid	fluent
	sBlum 2	3.3; +/-2.0	mid-soft	mid-fast	hyperfluent (1); disfluent (1)
	sSchoko 3	3.4; +/-2.1	mid	fast	fluent
	sSchoko 4	3.6; +/-2.1	mid	mid-fast	fluent

You

Deference	Speech sample	score	Overall loudness	Overall tempo	Fluency
	yLong 3	2.4; +1.3	mid-soft	slow	disfluent (1); hyperfluent (1)
	yLu3	2.7; +1.8	mid	slow	disfluent (1); hyperfluent (1)
	yLizhi 4	2.7; +1.9	mid	slow	disfluent (2)
	yMing 3	3.1; +1.5	mid	slow	disfluent (1)
Arrogance	Speech sample	score	Overall loudness	Overall tempo	Fluency
	yLizhi 2	4.6; - 2.4	loud	fast	hyperfluent (1)
	yLa 2	4.4; - 2.5	mid-loud	fast	disfluent (1); hyperfluent (2)
	yLong 2	4.3; - 2.3	mid-loud	mid-fast	fluent
	yMan 2	4.1; - 2.3	mid	fast	fluent
	yLu 2	3.6; - 2.1	mid	mid	disfluent (3); hyperfluent (1)
	yLong 1	3.6; - 2.1	mid	mid	fluent
	yLing 2	3.6; - 1.9	mid	mid-fast	disfluent (2)
	yDang 1	3.3; - 2.1	mid-loud	fast	hyperfluent (1)
Support	Speech sample	score	Overall loudness	Overall tempo	Fluency
	yLizhi 1	3.3; +2.2	mid	mid	fluent
	yLong 4	3.3; +2.4	loud	mid-fast	hyperfluent (2)
	yMan 1	3.1; +2.4	mid	mid-fast	hyperfluent (1)
	yLing 1	3.2; +2.1	mid	mid	fluent
	yMan 3	3.1; +1.9	mid-soft	slow	disfluent (5); hyperfluent (1.5)
	yBing 4	2.9; +1.9	mid	mid	fluent
Ambiguous	Speech sample	score	Overall loudness	Overall tempo	Fluency
	yMing 1	3.3; +/- 2.0	mid	mid-fast	hyperfluent (1)

Wu

Deference	Speech sample	score	Overall loudness	Overall tempo	Fluency
	wNing 4	1.8; +1.8	soft	mid-slow	fluent
	wLizhi 4	2.3; +1.4	soft	mid-slow	fluent
	wBing 3	3.0; +1.4	soft	mid-fast	fluent
	wLa 4	2.7; +1.9	mid	mid-slow	hyperfluent (1) disfluent (1)
	wLizhi 3	2.6; +1.8	mid-soft	mid	disfluent (1)
Arrogance	Speech sample	score	Overall loudness	Overall tempo	Fluency
	wLizhi 2	4.8; -1.8	loud	fast	hyperfluent (1)
	wBing 2	4.5; -2.1	loud	fast	hyperfluent (1)
	wMan 2	3.8; -1.9	mid	fast	hyperfluent (2)
	wLa 3	3.3; -1.9	mid	mid-fast	disfluent (1/2*)
Support	Speech sample	score	Overall loudness	Overall tempo	Fluency
	wDan 12	3.4; +2.1	loud	mid	fluent
	wDan 4	3.3; +2.2	mid	mid	fluent
	wMan 4	3.7; +1.7	mid	mid	hyperfluent (1/2)
	wLizhi 1	3.0; +2.4	mid	mid-slow	disfluent (2)
	wBing 1	3.0; +2.1	mid	mixed	hyperfluent (1) hyperfluent (1/2*) disfluent (2)
Ambiguous	Speech sample	score	Overall loudness	Overall tempo	Fluency
	wDan 11	3.2; +/- 2.4	mid	mid	fluent
	wMan 3	2.8; +/- 2.1	mid	mid-slow	disfluent (1)
	wLu 4	2.6; +/- 2.0	soft	mid	fluent
	wNing 2	2.4; +/- 1.3	loud	mid	hyperfluent (1)

Table 13 Mean tonal band-width, amplitude and emotional intensity

Speech sample	score	Value for mean tonal band-width (Hz)	amplitude (dB)	emotional intensity
yLong 3	2.4; +1.3	82	mid-soft	slightly elevated
yLu3	2.7; +1.8	83.5	mid	slightly elevated
yLizhi 4	2.7; +1.9	106.5	mid	strongly elevated
yMing 3	3.1; +1.5	≥ 122.6	mid	slightly elevated
yMan 2	4.1; - 2.3	≥ 42.5	mid	mid
yLizhi 2	4.6; - 2.4	136.5	loud	strongly elevated
yLong 2	4.3; - 2.3	57.3	mid-loud	strongly elevated
yLa 2	4.4; - 2.5	61.5	mid-loud	strongly elevated
yLu 2	3.6; - 2.1	≥ 68	mid	mid
yLing 2	3.6; - 1.9	75.6	mid	slightly elevated
yLong 1	3.6; - 2.1	110.8	mid	slightly elevated
yDang 1	3.3; - 2.1	119.5	mid-loud	strongly elevated
yBing 4	2.9; +1.9	≥ 60.6	mid	mid
yMan 3	3.1; +1.9	66.3	mid-soft	mid
yLing 1	3.2; +2.1	≥ 97.3	mid	slightly elevated
yLizhi 1	3.3; +2.2	97	mid	slightly elevated
yLong 4	3.3; +2.4	75.4	loud	strongly elevated
yMan 1	3.1; +2.4	≥ 78	mid	strongly elevated
yMing 1	3.3; +/-2.0	97.8	mid	strongly elevated

Table 13 You

Speech sample	score	Value for mean tonal band-width (Hz)	amplitude (dB)	emotional intensity
wNing 4	1.8; +1.8	77.3	soft	mid
wLizhi 4	2.3; +1.4	≥ 46	soft	low
wLa 4	2.7; +1.9	45.3	mid	low
wBing 3	3.0; +1.4	≥ 64	mid-soft	mid
wLizhi 3	2.6; +1.8	≥ 94.5	mid-soft	slightly elevated
wBing 1	3.0; +2.1	76.2	mid	mid
wLizhi 1	3.0; +2.4	106.9	mid	mid
wDan 4	3.3; +2.2	121.6	mid	slightly elevated
wMan 4	3.7; +1.7	167.5	mid	strongly elevated
wDan 12	3.4; +2.1	≥ 110.6	loud	strongly elevated
wBing 2	4.5; - 2.1	≥ 143	loud	strongly elevated
wMan 2	3.8; - 1.9	91 (?)	mid	strongly elevated
wLizhi 2	4.8; - 1.8	196.8	loud	strongly elevated
wLa 3	3.3; - 1.9	65	mid	mid
wDan 11	3.2; +/-2.4	187	mid	slightly elevated
wMan 3	2.8; +/-2.1	> 212.3	mid	strongly elevated
wLu 4	2.6; +/- 2.0	114.3	soft	mid
wNing 2	2.4; +/-1.3	199	mid	strongly elevated

Table 13 Wu

Table 14 Declination and syntactic form/illocutionary function

Declarative (statement)		Interrogative (without question particle)		Exclamative (exclamation)	
yLong 3	9.5	yLu 2	24.0	yLong 4	3.8
yLizhi 4	12.5	yMan 2	19.75	yLizhi 2	0.5
yMing 3	11.0	yLizhi 1	26.0		
yLong 1	7.5	yBing 4	≥ 22.0		
		yLong 2	18.0		
		yLa 2	18.0		
		yLing 2	20.0		
		yMing 1	20.5		
Mean:	10.1	Mean:	≥ 21.0	Mean:	2.15

Table 14 You

Declarative (statement)		Interrogative (with question particle)		Exclamative (exclamation)	
wNing 4	10.0	wLa 3	37.5	--	
wLizhi 3	7.25				
wLa 4	10.0				
wBing 3	12.5				
wMan 2	13.6				
wDan 4	11.6				
wLizhi 1	14.6				
wBing 1	10.5				
Mean:	11.25				

Table 14 Wu

Table 15 Final tonal behavior

Speech sample	score	Caudal tonal behavior (final tone/particle)	Focal tonal behavior
yLong 3	2.4; +1.3	2.5 tones (rise)	≥ 4 tones (fall); 2.5 tones (rise)
yLu3	2.7; +1.8	≥ 1 tone (rise); ≥ 1 tone (slope)	≥ 3 tone (rise); ≤ 3.5 tones (slope); ≥ 1 tone (rise); 1 tone (slope)
yLizhi 4	2.7; +1.9	≤ 2 tones (fall)	≤ 6 tones (fall); ≤ 2 tones (fall); together: ≤ 8 tones (overall fall)
yMing 3	3.1; +1.5	≤ 5.5 tones (slope)	≥ 5 tones (fall) ; ≤ 5.5 tones (rise); ≤ 5.5 tones (slope)
yLizhi 2	4.6; - 2.4	≤ 2.5 tones (fall)	5 tones (overall fall)
yLa 2	4.4; - 2.5	≤ 1.5 tones (fall)	≥ 2 tones (rise); ≥ 3 tones (overall fall)
yLong 2	4.3; - 2.3	≤ 1.5 tones (rise)	≥ 3 tones (fall); ≤ 1.5 tones (rise)
yMan 2	4.1; - 2.3	≥ 3 tones (fall)	≥ 3 tones (fall)
yLu 2	3.6; - 2.1	0.5 tone (fall)	≥ 3.5 tones (overall fall)
yLong 1	3.6; - 2.1	≤ 4.5 tones (slope)	5.5 tones (fall); ≤ 4 tones (rise); ≤ 4.5 tones (slope)
yLing 2	3.6; - 1.9	2.5 tones (fall)	≤ 3 tones (overall fall); ≥ 2.5 tones (overall fall)
yDang 1	3.3; - 2.1	≤ 5.5 tones (slope)	≤ 5.5 tones (slope)
yMing 1	3.3; +/-2.0	≥ 4 tones (rise)	2.5 tones (fall); ≥ 4 tones (rise); together: ≤ 5 tones (overall rise)
yLong 4	3.3; +2.2	3/5 tone (rise)	≤ 5 tones (fall); $\frac{3}{4}$ tone (rise); 3/5 tone (rise); in all: 1.5 tones (overall rise)
yMan 1	3.1; +2.4	$\frac{1}{4}$ tone (rise)	≤ 2.5 tones (fall); 2.5 tones (rise); $\frac{1}{4}$ tone (rise); in all: ≤ 3 tones (overall rise)
yLizhi 1	3.3; +2.2	≤ 1.5 tones (rise); $\frac{1}{5}$ tone (slope)	≥ 4.5 tones (fall); ≤ 1.5 tones (rise); $\frac{1}{5}$ tone (slope)
yLing 1	3.2; +2.1	6.5 tones (slope)	≤ 3.5 tones (fall); 6.5 tones (rise); 6.5 tones (slope)
yMan 3	3.1; +1.9	≤ 1 tone (fall)	≤ 4 tones (rise); 3.5 tones (fall); ≤ 1 tone (overall fall)
yBing 4	2.9; +1.9	≥ 1 tone (fall)	≥ 2 tones (overall fall)

Table 15 You

Speech sample	score	Tonal behavior of final tone/particle	Tonal behavior of final tone-unit
wNing 4	1.8; +1.8	≤ 1 tone (rise)	3.5 tones (fall); ≤ 1 tone (rise)
wLizhi 4	2.3; +1.4	1 tone (slope); 2 tones (rise)	≤ 2 tones (fall); 2 tones (rise); 1 tone (slope)
wLa 4	2.7; +1.9	~	≥ 2 tones (fall); \emptyset
wBing 3	3.0; +1.4	1.5 tones (fall)	4.5 tones (fall); 1.5 tones (fall)
wLizhi 3	2.6; +1.8	2 tones (slope)	≤ 2 tones (fall); 2 tones (slope); together: ≤ 4 tones (overall fall)
wLizhi 2	4.8; - 1.8	4 tones (rise)	≤ 11 tones (fall); ≤ 1.5 tones (fall); ≈ 12 tones (overall fall); 4 tones (rise)
wBing 2	4.5; - 2.1	≤ 2 tones (rise)	≥ 4.5 tones (fall); ≥ 2 tones (fall); ≥ 1 tone (fall); = 7.5 tones (overall fall); ≤ 2 tones (rise)
wMan 2	3.8; - 1.9	≥ 2 tones (slope)	≤ 8.5 tones (fall); ≥ 2 tones (slope); in all: ≥ 10 tones (overall fall)
wLa 3	3.3; - 1.9	≤ 2 tones (fall)	5.5 tones (rise); ≤ 2 tones (fall)
wDan 12	3.4; +2.1	\emptyset	≥ 4.5 tones (fall)
wDan 4	3.3; +2.2	$\geq 3/4$ tone (slope)	5 tones (fall); $\leq 1/2$ tone (rise); $3/4$ tone (slope)
wMan 4	3.7; +1.7	≥ 1.5 tones (fall)	≥ 2 tones (rise); 1.5 tones (fall)
wLizhi 1	3.0; +2.4	≤ 2.5 tones (fall)	2.5 tones (fall)
wBing 1	3.0; +2.1	≥ 3 tones (fall)	3 tones (rise); ≥ 3 tones (fall)
wDan 11	3.2; +/-2.4	≤ 2 tones (fall)	≤ 10 tones (fall); ≤ 9 tones (rise); 2 tones (fall)
wMan 3	2.8; +/-2.1	10 tones (rise); ≤ 9.5 tones (slope)	≥ 10 tones (fall); 10 tones (rise); ≤ 9.5 tones (slope)
wLu 4	2.6; +/-2.0	1.5 tones (fall)	≤ 9 tones (rise); 7 tones (fall); 1.5 tones (fall)
wNing 2	2.4; +/-1.3	≤ 3 tones & ≤ 5.5 tones (rise-fall)	≤ 8 tones (slope); ≤ 3 tones & ≤ 5.5 tones (rise-fall)

Table 15 Wu

10. Final conclusion: The signaling of interpersonal attitude in German and Chinese

According to the findings of interactional sociolinguistics, the communicative phenomena of greatest import to intercultural communication are the suprasegmental features, whose use in the signaling of referential, social and attitudinal meaning is regulated by socio-culturally patterned rules and often varies strongly from one speech community to the next. Given the importance of the attitudinal dimension of communication for human relationships, one of the most important research topics in intercultural communication is the manner in which different speech communities communicate interpersonal attitude, relating to *footing*, the power position between the speaker and the listener, and *affinity* the degree to which the speaker identifies himself with his interaction partner. As a contribution to intercultural communication research involving the Chinese language, the experimental study undertaken within the scope of this work examined the manner in which speakers of German and Chinese use the suprasegmental features - intonation, overall pitch, loudness, tempo and fluency, and voice quality - to signal different interpersonal attitudes, defined in terms of the interaction strategies *camaraderie*, *distance*, *support* and *deference* and their antagonists *pushiness*, *cold aloofness*, *arrogance* and *false deference*¹. The experimental study focused on four questions:

- (1) Are the interaction strategies *camaraderie*, *support*, *distance*, *deference*, *arrogance* and *false deference* found in German and Chinese conversational interactions?
- (2) Does the ambiguity of these interaction strategies, which was discussed above for the English language, also apply to German and Chinese?
- (3) In what way do the three groups of suprasegmental phenomena - intonation, the overall features of tempo, loudness, pitch and fluency, and voice quality - contribute to the signaling of *support*, *deference*, *arrogance* and *false deference* in German and Chinese?
- (4) Are the suprasegmental features contributing to the signaling of these interaction strategies responsible for their ambiguity and the misunderstandings of speaker attitude in German - Chinese intercultural communication?

In the following, the main results of the experimental study and their implications for intercultural communication are discussed separately for each of these objectives.

10.1 Interaction strategies in German and Chinese conversational interaction

With respect to questions (1) and (2), pertaining to the use and ambiguity of the above named interaction strategies in the two languages under discussion, the results of the macroanalysis of the German and Chinese dialogs (Chapter 6) showed that *camaraderie*, *support*, *arrogance*, *distance* and *deference* are indeed part of conversational interaction in German and Chinese society. This was confirmed once more for *support*, *arrogance* and *deference* by the results of the sociolinguistic judgement tests (Chapter 7), according to which 74 (71%) of 103 German and Chinese utterances could be classified in terms of one, maximally two, of these interaction strategies by native speakers. On the other hand, the results of these tests also showed that the communication and interpretation of these attitudinal messages involves the same difficulties as have been described for English: In addition to the 29 utterances which could not be analyzed at all in terms of speaker attitude, 16 speech samples were judged to be ambiguous with respect to the *valence*, or undertone, signaled therein. Taken together, this gives a total of 45 German and Chinese speech samples (43.7% of the corpus) which posed moderate to strong difficulties in the determination of interpersonal attitude. Further evidence for the ambiguity of the interaction strategies and their attitudinal metamessages was obtained

¹ It will be recalled from Part One that the antagonists of *camaraderie*, *support*, *distance* and *deference* are the interpersonal attitudes, or interaction strategies, in terms of which they are most commonly misunderstood (See summary of Part One). The numerous ways in which *deference* can be misunderstood are taken together here under the term *false deference*.

in the second part of the judgement tests, in which the large majority of the German and Chinese speech samples were judged differently than in the first part. These findings indicate that *support*, *arrogance* and *deference* are as well-known in Chinese and German society as in English speech communities and that they also involve the same difficulties as discussed for the English language.

As the third result relating to questions (1) and (2), evidence for a number of differences between German and Chinese were also found with respect to the use of the interaction strategies: To begin with, two of the German utterances with ambiguous *valence* were describable in terms of ambiguous, and therefore potentially *false*, *deference*, but none of the Chinese speech samples was classified in terms of this interaction strategy. This was interpreted as evidence for the strong conventionalization of the *deference* strategy in Chinese society. As a second difference, two speech samples in the Chinese corpus were found to represent the challenging, teasing humour strategy of *tai-gang*. No evidence of such teasing was found in the German utterances, pointing to differences in the use of irony and teasing in Chinese and German society. It appears, for example, that in contrast to China, where *tai-gang* may be encountered in conversations between people meeting for the first time, the use of irony and teasing in German society is largely limited to interactions between friends or at least acquaintances. As a third difference between the two languages, one Chinese speech sample was found to represent the Chinese strategy of *sa-jiao*, a frequently encountered feature of women's speech, in which the speaker, with a wailing, whimpering tone of voice and a highly exaggerated manner of speaking, tries to get her way with the listener.

10.2 The signaling of interpersonal attitude in German and Chinese

Coming now to question (3), the results of the microanalyses of Chapter 8 and, in particular, the comparative examinations of the German and Chinese suprasegmental profiles in Chapter 9, yielding the phonetic exponency of *deference*, *support* and *arrogance* in the two languages, revealed that the communication of these three strategies is achieved in much the same way in German and Chinese, involving largely the same suprasegmental features.

Thus, in both languages *deference* is communicated by (*modal-*)*lax* and/or *breathy* phonatory voice quality, *tenser* phonatory *settings* found in utterances signaling a strong *footing*. The use of the supralaryngeal *setting* of *palatalization*, optionally accompanied in Chinese by delayed *glottalization*, was also observed in both languages. *Deference* is further signaled by mid to reduced, only rarely elevated, loudness, tempo and fluency. Reduced values for more than one of these phenomena signal a weak, inferior *footing*, while elevated values are common in utterances communicating a stronger, more forceful attitude. Mid to high overall pitch is also typical of deferent utterances in both languages. This feature is linked to the signaling of weak *footing* in German, to the communication of positive *valence* in Chinese. With regards to intonation, the weak, inferior *footing* typical of *deference* is communicated in German by narrow, and often high, nuclear pitch-patterns, in Chinese by weak, narrow, focal tonal behavior. A stronger, more forceful attitude is signaled by wider nuclear pitch-patterns and wider focal tonal behavior. A further feature of *deferent* utterances in German is their relatively narrow prenuclear pitch-range, linked to the positive *valence*/undertone of this strategy. Coming to the results for *support*, in German and Chinese this strategy is communicated by *modal-laxness*, *lax-* and/or *breathiness*, the latter more frequent in Chinese than in German, where *breathiness* appears to be more typical of *deference*. In Chinese, utterances often begin with a relatively *tense* phonatory *setting*, which then develops into *laxer*, softer, often *breathy* voice quality by mid-utterance. *Palatalization*, frequent among the deferent utterances, was not found in any supportive speech samples in the corpus, a finding which suggests that it can be used to distinguish between *deference* and *support* in both languages. *Support* is further signaled by mid tempo, loudness and fluency, elevated values for one or more of these

phenomena found in utterances signaling a strong *footing*, while reduced values are characteristic of utterances signaling a weaker, inferior, attitude on the part of the speaker. Conversely, a weak *footing* is signaled in German by a high, elevated, overall pitch, a lower pitch being more common in the communication of a strong, forceful attitude. Given its link to positive *valence* in Chinese, overall pitch is generally relatively high in Chinese supportive utterances. With regards to intonation, nuclear pitch-patterns and focal tonal behavior are wide and strong in supportive utterances signaling a strong, superior *footing*, less wide in utterances communicating a gentler, less forceful attitude. In German, an additional feature of *support* is relatively narrow prenuclear pitch-range linked to the positive *valence*/undertone of this strategy. Finally, in both languages *arrogance* is signaled by the pronounced presence of *modalness*, blending frequently with *modal-tenseness* and/or *tenseness*, only rarely with *laxness* or *modal-laxness*. *Tenseness* on its own is found in strongly arrogant utterances, while *laxer*, softer phonatory *settings* are common in less strongly arrogant utterances. *Palatalization* is not employed in this interaction strategy. In both languages *arrogance* is further signaled by mid to elevated loudness, tempo and fluency. Elevated values for more than one of these phenomena communicate strong *arrogance*, while mid, and occasionally also reduced, values for the overall features are common in utterances signaling a weaker, less forceful *footing*, or attitude. In both languages, an additional feature of *arrogance* is a relatively low overall pitch, linked in German to the signaling of a strong, forceful attitude, in Chinese to the negative *valence*/undertone characteristic of this strategy. With regards to intonation, the strong, superior *footing* typical of arrogant utterances is signaled in German by wide nuclear pitch-patterns and in Chinese by wide focal tonal behavior, weaker, i.e. narrower, nuclear pitch-patterns and focal tonal behavior used to communicate a weaker, less forceful speaker attitude. Finally, the negative undertone of *arrogance* is also communicated in German by a relatively wide prenuclear pitch-range.

With respect to question (4), pertaining to the precise role of the suprasegmental features in these communicative processes, the findings of Chapter 9 indicate that the contribution of these features is also largely the same in the two languages. As a first point, the phonatory *settings* and the overall features contribute in almost identical manner to the signaling of *valence*/undertone and the (relative) *footing* of utterances, the only marked difference between the two languages being found with respect to the use of *breathiness* and overall pitch. Whereas *breathiness* is a feature of both *support* and *deference* in Chinese, in German it is limited to the strategy of *deference*. And while overall pitch is linked to *footing* in German, it is associated with *valence*/undertone in Chinese. As a second point, the supra-laryngeal *settings* of *palatalization* and *labiodentalization* were also found to hold similar functions in German and Chinese - the signaling of relative weakness and in the latter case potentially also positive *valence*/undertone. As *palatalization* is found most frequently in the deferent utterances of the speech corpus, it was concluded that it may also be regarded as a marker of the *deferent* strategy, again in both languages. Finally, turning to the contribution of intonation, though the examinations did not include Chinese counterparts of German and English *heads*, the results of the above analyses showed that the nuclear pitch-patterns of German (and English) intonation and their Chinese correlates, the focal tonal behavior of Mandarin Chinese utterances, hold the same attitudinal functions, the signaling of the relative strength, or *footing*. Evidence for a functional similarity was also found between high, narrow leveling-off pitch-patterns in German and narrow leveling-off caudal tonal behavior in Chinese, both of which signal a weak, inferior *footing*, and between German wide low rises and Chinese low rising caudal behavior, both signaling a strong, potentially negative attitudinal message.

On the basis of these results, it may be concluded that the role of the suprasegmental features in the signaling of *valence* and *footing*, and therefore also in the communication of the

interaction strategies under discussion here, is largely the same in the two languages, the only differences found with respect to the use of *breathiness* and overall pitch. This in turn suggests that if misunderstandings of speaker attitude occur in German - Chinese intercultural interaction these can only be partly blamed on differences with respect to the use of the suprasegmental features, since only *breathiness* and overall pitch are used in a slightly different way.

10.3 Implications of the obtained results for intercultural communication

At first sight, the above discussed results, according to which the signaling of interpersonal attitude in German and Chinese involves largely the same signaling processes based on the same suprasegmental features, suggest that the communication of attitude in cross-cultural encounters between speakers of these two languages should be largely unproblematic, involving only differences in politeness norms, relating to the different habitual use of the interaction strategies in social intercourse. However, while not wanting to belittle the problems which can occur on this interactional level of communication², a closer look at the results obtained in Parts Two and Three show that such a view is rather too optimistic for a number of reasons. First of all, it must be recalled that, due to the multifunctional and therefore highly ambiguous nature of the suprasegmental features, the communicative patterns set up in the course of Chapter 9 do not function independently of each other, so that the precise communicative effect of any one phonatory or supralaryngeal *setting*, overall feature and form of intonation in a given utterance - their attitudinal message - depends to no small degree on that of other features present in the utterance. As one example, it was found that German utterances in which the *lax*, gentle and weak phonatory *settings* and overall features typical of *deference* are found together with strong, potentially negative intonation features, such as 'drop' *nuclei* and *glissando*, signal a highly ambiguous attitudinal message.

As a second caveat, relating to intonation, it must be pointed out that the comparative examinations undertaken here were limited to one sole intonational form, utterance-final intonational pitch-patterns. No effort was undertaken, indeed could be undertaken at this stage of research, to examine the existence and functions of 'prenuclear' pitch-patterns, i.e. 'pre-focal' tonal behavior, in Chinese. It is not inconceivable that future research will reveal differences here across the two languages - potential sources of miscommunication. Also, strictly speaking, the discovered similarities relate largely to the phenomenon of pitch-range and only two concrete intonational pitch-patterns, high, narrow falls and low wide rises, could be compared with each other. Further research into Chinese *caudal tonal behavior* will show whether more similarities exist or whether these are the only common points between two quite different systems. Finally, with respect to the production of intonational forms in the two languages, it must be pointed out that the primary role of *pitch-range*, in effect *tonal bandwidth*, not only in the creation of Chinese intonational contours but also in the signaling of the *focus*, poses immense difficulties to native speakers of western languages, such as English and German, accustomed to variations in *pitch*. Therefore, communication between native speakers and western learners of Chinese not yet proficient in the signaling of these *pitch-range patterns* is likely to be disturbed, involving much the same difficulties as have been described for other languages. Two further, related, differences were also found with respect to loudness, which in contrast to its use in western languages, does not contribute to the signaling of the *focus* in Chinese. The results of the study also suggest that, in difference to German, where this feature is used to signal emphasis as well as emotional involvement, only the second use of loudness applies to Chinese, a difference which also constitutes a potential source of miscommunication.

² These difficulties were discussed in detail in Chapter 1.

Third and lastly, it must be remembered that, according to the research results achieved in interactional sociolinguistics, the most frequent and most unhappy misunderstandings of attitudinal messages in intercultural communication originate not so much in misunderstandings of attitudinal signals but in misinterpretations of messages originally meant to signal purely linguistic, non-attitudinal, information. Since such misunderstandings, involving a confusion, in effect, of the linguistic and paralinguistic/extralinguistic dimensions of suprasegmental communication, are not caused by the communicative functions of the suprasegmental features but originate in their intrinsic functional diversity, and ambiguity, even perfect knowledge of these signaling processes cannot protect interactants from such instances of miscommunication.

Bibliography

- A New Practical Chinese-English Dictionary. 1972. Liang Shih-Chiu et al. (eds.). Taipei: Far East Book Company
- Abberton, E. & A. J. Fourcin. 1972. Laryngographic analysis and intonation. *British Journal of Disorders of Communication* 7, 24 - 29
- Abercrombie, David, D. B. Fry, P.A. D. McCarthy, N. C. Scott and J. L. M. Trim (eds.) 1964. In Honour of Daniel Jones: Papers Contributed on the Occasion of his Eightieth Birthday. London: Longman
- Abercrombie, David, 1967. *Elements of General Phonetics*. Edinburgh: Edinburgh University Press
- Addington, D. W. 1968. The relationship of selected vocal characteristics to personality perception. *Speech Monographs* 35, 492 - 503
- Allport, G. W. & H. Cantril. 1934. Judging personality from voice. *Journal of Social Psychology* 5, 37 - 54
- Anderson, Keith. O. 1978. On the contrastive phonetics of English and German intonation. In: *Festschrift für Otto von Essen anlässlich seines 80. Geburtstages*. *Hamburger Phonetische Beiträge* 25, 25 - 35. Hamburg: Buske
- Anderson, Stephen. 1985. *Phonology in the Twentieth Century. Theories of Rules and Theories of Representations*. Chicago and London: University of Chicago Press
- Apple, W., L. A. Streeter & R. M. Krauss. 1977. The effects of pitch and speech rate on personal attributions. MS. Columbia University
- Apple, W. L., L. A. Streeter & R. M. Krauss. 1979. Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology* 37, 715 - 727
- Aronovitch, C. D. 1976. The voice of personality: Stereotyped judgements and their relation to voice quality and sex of speaker. *Journal of Social Psychology* 99, 207 - 220
- Atkinson, J. M. & J. Heritage (eds.). 1984. *Structures of Social Action. Studies in Conversation Analysis*. Cambridge: Cambridge University Press
- Auer, Peter. 1986. Kontextualisierung. *Studium Linguistik* 19, 22 - 47
- Austin, John L. 1962. *How to do things with words*. Oxford: Clarendon Press
- Bailey, C.-J. 1983. A concise but comprehensive approach to intonation for learners of English. *Arbeiten aus Anglistik und Amerikanistik* 8, 3 - 27
- Baker, Hugh, D. R. 1979. *Chinese family and kinship*. New York
- Bald, W.-D. 1976. Contrastive studies in English and German intonation: a survey. In Fisiak (ed.)
- Banse, R. & Klaus R. Scherer. 1996. Acoustic Profiles in Vocal Emotion Expression. *Journal of Personality and Social Psychology* 70 (3), 614 - 636
- Barry, W. J. 1981. Prosodic functions revisited again! *Phonetica* 38, 320 - 340
- Bateson, Gregory. 1972. *Steps to an Ecology of Mind*. New York: Ballantine.
- Bauman, R. & J. Sherzer (eds.). 1974. *Explorations in the ethnography of speaking*. Cambridge: Cambridge University Press.
- de Beaugrande, Robert. 1983. Text production: towards a science of composition. *Advances in Discourse Processes* 11. Norwood, NJ: Ablex
- Bell, Alexander G. 1908. *The mechanisms of speech*. New York: Funk & Wagnalls
- Bennett, S. & B. Weinberg. 1979. Sexual characteristics of pre-adolescent children's voices. *Journal of the Acoustical Society of America* 65, 179 - 189
- van den Berg, J. 1958. Myoelastic - aerodynamic theory of voice production. *Journal of Speech and Hearing Research* 1, 227 - 244
- van den Berg, J. 1968. Mechanisms of the larynx and laryngeal vibrations. In: Malmberg (ed.), 278 - 308

- Bergman, Jörg. 1981. Ethnomethodologische Konversationsanalyse. Schröder, P. & H. Steger (eds.), 9 - 51
- Bergemann, Niels & A. Sorrisseaux (eds.) 1992. Interkulturelles Management. Heidelberg: Physica
- Blakar, Rolv M. 1975. An experimental method for enquiring into communication. *European Journal of social psychology* 3 (4), 415 - 425
- Blom, J. P. & Gumperz, J. J. 1972. Social meaning in Linguistic Structure: Code-switching in Norway. In Gumperz & Hymes (eds.), 407 - 434
- Bolinger, Dwight L. 1958. A theory of pitch-accent in English. *Word* 14, 109 - 149
- Bolinger, Dwight L. 1964. Intonation as a universal. In H. G. Lunt (ed.) *Proceedings of the 9th International Congress of Linguists, 1962*. The Hague: Mouton, 833 - 844
- Bolinger, Dwight L. (ed.). 1972a. *Intonation. Selected Readings*. Harmondsworth: Penguin
- Bolinger, Dwight L. (1972b). Accent is predictable (if you're a mind-reader). *Language* 48, 633 - 644
- Bolinger, Dwight L. 1978. Intonation accross languages. In J. H. Greenberg, C. A. Ferguson & E. A. Moravcsik (eds.), 471 - 524
- Bolinger, Dwight L. 1986. Intonation and emotion. *Quaderni di Semiotica* 7, 13 - 21
- Bolinger, Dwight L. 1989. *Intonation and its uses. Melody in Grammar and Discourse*. Stanford: Stanford University Press
- Bond, Michael H. 1996. *The Handbook of Chinese Psychology*. Oxford & New York: Oxford University Press
- Borden, Gloria J. & Katherine S. Harris. ²1984. *Speech Science Primer: Physiology, Acoustics and Perception of Speech*. Baltimore: Williams and Wilkins
- Broad, D. J. 1979. New theories of vocal fold vibration. In Lass (ed.), vol. 1, 203 - 256
- Brend, R. 1972. Male - female intonation patterns in American English. *Proceedings of the 7th International Congress of Phonetic Sciences*, Mouton: 866 - 869. Also in Thorne & Henley (eds.) 1975, 84 - 87
- Brown, B. L., W. J. Strong & A. C. Rencher. 1974. Fifty-four voices from two: the effects of simultaneous manipulations of rate, mean fundamental frequency, and variance of fundamental frequency on ratings of personality from speech. *Journal of the Acoustical Society of America* 55, 313 - 318
- Brown, B. L., W. J. Strong & A. C. Rencher. 1975. Acoustic determinants of the perceptions of personality from speech. *International Journal of the Sociology of Language* 6, 11 - 32
- Brown, Gillian & George Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press
- Brown, Penelope & Levinson, Stephen. 1978. Universals in language usage: Politeness phenomena. E. Goody (ed.), 58 - 289
- Brown, Penelope & Stephen C. Levinson. 1987. *Politeness. Some Universals in language usage*. Cambridge: Cambridge University Press (reprint)
- Brown, Roger & Gilman, Albert, 1960. The pronouns of power and solidarity, *American Anthropologist* 4 (6), 24 - 29
- Catford, J. C. 1964. Phonation types: The classification of some laryngeal components of speech production. in Abercrombie et al (eds.), 26 - 37
- Catford, J. C. 1977. *Fundamental Problems in Phonetics*. Edinburgh: Edinburgh University Press
- Chafe, Wallace L. 1974. Language and consciousness. *Language* 50, 111 - 133
- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Li (ed.), 25 - 55.
- Chambers, J. K. 1995. *Sociolinguistic Theory. Linguistic Variation and its Social Significance*: Oxford & Cambridge: Blackwell

- Chan, Marjorie. <http://deall.ohio.state.edu/chan.9/MToBI.htm>
- Chang, N. C. 1958. Tones and Intonation in the Chengtu dialect (Szechuan, China). *Phonetica* 1-2: 59 - 85. Reprinted in Bolinger (ed.) 1972a
- Chao, Yuen-Ren. 1933. Tone and Intonation in Chinese. *Bulletin of the Institute of History and Philology* 4, 121 - 134
- Chao, Yuen-Ren. 1948. *Mandarin Primer*. Cambridge, Ma.: Harvard University Press
- Chao, Yuen-Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press
- Cheshire, Jenny. 1982. Linguistic variation and social function. In Romaine (ed.), 153 - 166.
- Childers, D. G., A. M. Smith & G. P. Moore. 1984. Relationships between electroglottograph, speech and vocal cord contact. *Folia Phoniatrica* 36, 105
- Childers, D. G. & A. K. Krishnamurthy. 1985. A critical Review of electroglottography. *CRC Critical Reviews in Biomedical Engineering* 12 (2), 131 - 161
- Clark, John & Colin Yallop. 1990. *An Introduction to Phonetics and Phonology*. Oxford: Blackwell
- Clyne, Michael. 1996. *Intercultural Communication at work*. Cambridge: Cambridge University Press
- Coates, Jennifer. 1986. *Men, Women and Language*. London: Longman
- Coates, Jennifer (ed.) 1998. *Language and Gender: A Reader*. Oxford: Blackwell
- Coates, Jennifer & Deborah Cameron. 1989. (eds.) *Women in Their Speech Communities*. London: Longman
- Cohen, A. & S.G. Nooteboom. 1975. *Structure and Process in Speech Perception*. Berlin: Springer
- Cohen, A., R. Collier & J. 't Hart. 1982. Declination: Construct or intrinsic feature of speech? *Phonetica* 39, 254 - 273
- Cole, Peter & Jerry L. Morgan (eds.). 1975. *Syntax and Semantics*, vol. 3. New York: Academic Press
- Connell, B. & D. R. Ladd. 1990. Aspects of pitch realization in Yoruba. *Phonology* 7, 1 - 29.
- Coster, D.C. & Kratochvil, Paul, 1984. Tone and stress discrimination in normal Beijing dialect speech. *New papers on Chinese language use*, 119 - 132. Canberra
- Coulthard, Malcomb. 1977. *An Introduction to Discourse Analysis*. London: Longman
- Couper-Kuhlen, Elizabeth. 1983. Intonatorische Kohäsion: Eine makroprosodische Untersuchung. In Klein (ed.), 74 - 100
- Couper-Kuhlen, Elizabeth. 1986. *An Introduction to English Prosody*. Tübingen: Niemeyer
- Couper-Kuhlen, Elizabeth. 1988. Contextualizing discourse. The prosody of interactional repair. Paper held at the Conference 'Contextualization of Language', University of Konstanz
- Couper-Kuhlen, Elizabeth. 1993. *English Speech Rhythm: Form and Function in everyday verbal interaction*. Amsterdam: John Benjamins
- Couper-Kuhlen, Elizabeth. 1996. The prosody of repetition: On quoting and mimicry. In Couper-Kuhlen & Selting (eds.), 366 - 405
- Couper-Kuhlen, Elizabeth. 1997. Cheer up for Goodness Sake - Prosodische Aspekte der sozialen Konstruktion von 'Gender'. In *Gender Studies an der Universität Konstanz*. 1996/1997. Frauenrat der Universität Konstanz. April 1997, 168 - 181
- Couper-Kuhlen, Elizabeth & M. Selting (eds.) 1996. *Prosody in Conversation. Interactional Studies*. Cambridge: Cambridge University Press
- Cruttendon, Alan. 1986. *Intonation*. Cambridge: Cambridge University Press
- Crystal, David. 1969. *Prosodic Systems and Intonation in English*. Cambridge: Cambridge University Press

- Crystal, David. 1975. *The English Tone of Voice. Essays in intonation, prosody and para-language*. London: Edward Arnold
- Crystal, David. 1985. *A Dictionary of Linguistics and Phonetics*. Oxford: Basil Blackwell
- Crystal, David, 1987. *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press
- Currie, Karen L. 1979. Contour systems of one variety of Scottish English. *Language and Speech* 22, 1 - 20
- Daniloff, R. (ed.) 1985. *Speech science*. London: Taylor and Francis
- Darby, J. K. (ed.) 1981. *Speech evaluation in psychiatry*. New York: Grune & Stratton
- Darwin, Charles. 1872. *The Expression of the Emotions in Man and Animals*. London: Murray
- Darwin, Charles, 1874. *The descent of Man and Selection in Relation to Sex*. Detroit: Gale Research Company
- Davitz, J. R. (ed.) 1964a. *The Communication of Emotional Meaning*. New York: McGraw-Hill
- Davitz, J. R. 1964b. A Review of Research Concerned with Facial and Vocal Expression of Emotion. In Davitz (ed.), 13 - 30
- Davitz, J. R. 1964c. Auditory Correlates of Vocal Expressions of Emotional Feeling. In Davitz (ed.), 101 - 112
- Di Cristo, Albert. 1998. Intonation in French. In Hirst & Di Cristo (eds.), 195 - 218
- Diehl, C. F., R. White & K. W. Burk. 1959. Voice quality and anxiety. *Journal of Speech and Hearing Research* 2, 282 - 285
- Duden Aussprachewörterbuch. 1990. Band 6. Mannheim, Wien & Zürich: Dudenverlag
- Duranti, A. and D. Brenneis (eds.) 1986. The audience as co-author. Special issue of *Text* 6 (3), 239 - 247
- Durkheim, E. 1915. *The elementary forms of the religious life*. New York: The Free Press.
- Duncan, S. 1973. Towards a Grammar for Dyadic Conversation. *Semiotica* 9, 29 - 46
- Duncan, S. 1974. On the Structure of Speaker-Auditor Interaction during Speaking Turns. *Language in Society* 3, 161-180
- Edelsky, Carole. 1979. Question intonation and sex roles. *Language in Society* 8, 15 - 32
- Ekman, P. W., R. E. Sorensen & W. V. Friesen. 1969. Pan-cultural elements in facial displays of emotion *Science* 164, 86 - 88
- Ekman, P. W. V. Friesen & K. R. Scherer. 1976. Body movement and voice pitch in deceptive interaction. *Semiotica*, 16, 23 - 27
- Elyan, O., P. Smith, H. Giles & R. Bourhis. 1978. RP - accented female speech: The voice of perceived androgyny? In Trudgill (ed.), 122 - 182
- Enkvist, Nils E, H. Ringbom et al. (eds.). 1975. *Style and Text*. Stockholm: Skriptor
- Erickson, F. & J. Schultz. 1982. *The counselor as gatekeeper: Social interaction in interviews*. New York: Academic Press
- Ervin-Tripp, S. 1964. An analysis of the interaction of language, topic and listener. *American Anthropologist* 66 (6; part 2), 86 - 102
- von Essen, Otto. 1964. *Grundzüge der Hochdeutschen Satzintonation*. Ratingen: Henn
- Esser, Jürgen. 1978. Contrastive intonation of German and English. Problems and some results. *Phonetica* 35, 41 - 55
- Fairbanks, John K. & Edwin O. Reischauer. 1973. *China: Tradition and Transformation*. Houghton Mifflin
- Fairbanks, G. & W. Pronovost. 1939. An Experimental Study of the Pitch Characteristics of the Voice during the Expression of Emotion. *Speech Monograph* 6, 87 - 104
- Fasold, Ralph. 1990. *The Sociolinguistics of Language*. Oxford: Basil Blackwell

- Fillmore, Charles J. 1974. Pragmatics and the description of discourse. In Schmidt (ed.), 83 - 104
- Fillmore, Charles J. 1976. The need for frame semantics within linguistics. *Statistical methods in linguistics*: 5-29. Stockholm: Skriptor
- Fillmore, Charles J. 1982. Frame semantics. *Linguistic Society of Korea. Linguistics in the Morning Calm*: 111 - 138. Seoul, Korea
- Fisiak, Jacek. (ed.) 1976. *Papers and Studies in Contrastive Linguistics* 4. Poznań
- Fónagy, I. 1981. Emotions, Voice and Music. In Sundberg, J. (ed.), 51 - 79
- Fónagy, I. & K. Magdics. 1963. Emotional Patterns in Intonation and Music. *Zur Phonetischen und Sprachwissenschaftlichen Kommunikationsforschung* 16, 293 - 326
- Fónagy, I. & E. Bérard. 1972. Il est huit heures: contribution à l'analyse sémantique de la vive voix. *Phonetica* 26, 157 - 192
- Fourcin, A. J. 1981. Laryngographic assessment of phonatory function. *ASHA Reports (American Speech-Language-Hearing Association)* 11, 117 - 127
- Fourcin, A. J. & E. Abberton. 1971. First Applications of a New Laryngograph. *Medical and Biological Illustrations* 27, 172 - 183
- Fourcin, A. J. & E. Abberton. 1977. Laryngograph studies of vocal fold vibration. *Phonetica* 34, 313
- Fox, Anthony. 1984. *German Intonation. An outline*. Oxford: Clarendon Press
- Freedman, Maurice (ed.) 1970. *Family and Kinship in Chinese Society*. Stanford: Stanford University Press
- Frick, R. W. 1985. Communicating Emotion: The Role of Prosodic Features. *Psychological Bulletin*. 97, 412 - 429
- Fromkin, Victoria A. (ed.) 1978. *Tone: A Linguistic Survey*. New York: Academic Press.
- Gårding, Eva. 1987. Speech Act and tonal pattern in standard Chinese, constancy and variation. *Phonetica* 44: 13 -29
- Garfinkel, Harold. 1967. *Studies in ethnomethodology*. Englewood Cliffs, NJ.: Prentice Hall
- Garfinkel, Harold. 1972. Remarks on Ethnomethodology. In Gumperz & Hymes (eds.), 301 - 324
- Gao, N. 1983. Hauptschwierigkeiten für Chinesen beim Erlernen der Deutschen Sprache. *Informationen Deutsch als Fremdsprache* 3, 41 - 48
- Gibbon, Dafydd. 1998. Intonation in German. In Hirst & Di Cristo (eds.), 78 - 95
- Gibbs, R. W. 1986. On the Psycholinguistics of Sarcasm. *Journal of Experimental Psychology: General* 115, 3 - 15
- Giglioli, Pier P. (ed.) 1972. *Language and Social Context*. Harmondsworth: Penguin
- Giles, H. 1970. Evaluative reactions to accents. *Educational Review* 22, 211 - 227
- Giles, H. 1971. Patterns of evaluation in reactions to RP. South Welsh and Somerset accented speech. *British Journal of Social and Clinical Psychology* 10, 280 - 281
- Giles, H. & P. F. Powesland. 1975. *Speech style and social evaluation*. London: Academic Press
- Giles, H. & R. Y. Bourhis. 1973. Dialect perception revisited. *Quarterly Journal of Speech* 59, 337 - 342
- Giles, H. & P. Marsh. 1979. Perceived Masculinity, Androgyny and Accented Speech. *Language Sciences* 1, 301- 315
- Giles, H., P. Smith, B. Ford, S. Condor & J. Thakerer (eds.). 1980. Speech style and the fluctuating saliency of sex. *Language Sciences* 2, 260 - 282
- Giles, H., W. P. Robinson & P. Smith. (eds.). 1980a. *Language: social psychological perspectives*. Oxford: Pergamon
- Giglioli, P. (ed.) 1972. *Language and social context*. Harmondsworth: Penguin
- Gimson, A.C.⁴1989. *An introduction to the Pronunciation of English*. London: Edward Arnold

- Gladwin, T. & W.C. Sturtevant (eds.). 1962. *Anthropology and Human Behavior*. Washington DC: Anthropology Society of Washington
- Goffman, Erving. 1961. *Encounters*. Indianapolis: Bobbs Merrill
- Goffman, Erving. 1967. *Interaction Ritual*. Garden City, N.J.: Anchor Books
- Goffman, Erving. 1974. *Frame Analysis*. New York: Harper and Row
- Goldman-Eisler, Frieda. 1968. *Psycholinguistic expression in spontaneous speech*. London: Academic press
- Goody, Esther (ed.). 1978. *Questions and Politeness*. Cambridge: Cambridge University Press
- Graddol, David & Swann, Joan. 1989. *Gender Voices*. Oxford: Basil Blackwell
- Granet, Marcel. 1958. *Chinese Civilization*. Cleveland and New York: Meridian
- Greenbaum, Sidney & Quirk, R. 1990. *A Student's Grammar of the English Language*. London: Longman
- Greenberg, J. H., C. A. Ferguson & E. A. Moravcsik (eds.). 1978. *Universals of Human Language*, vol 2: Phonology. Stanford: Stanford University Press
- Grice, Horace P. 1957. *Meaning*. *Philosophical Review* 66, 377 - 388
- Grice, Horace P. 1975. *Logic and Conversation*. In Cole & Morgan (eds.), 41 - 58
- Grice, Martine & R. Benz Müller. 1995. *Transcription of German Intonation using ToBI-Tones - The Saarbrücken System*. Saarbrücken: Phonus 1, 33 - 51
- Grice, Martine & Stefan Baumann. 2000. *Deutsche Intonation und GTobI*. *Linguistische Berichte* 181. Hamburg: Buske, 1 - 33
- Gu, Yue-Guo. 1990. *Politeness Phenomena in Modern Chinese*. *Journal of Pragmatics* 14, 237 - 257
- Gudykunst, William B., T. Nishida & S. Ting-Toomey. 1996. *Communication in Personal Relationships Across Cultures*. Thousand Oaks: Sage
- Gumperz, John J. 1971. *Language in Social Groups*. Stanford: Stanford University Press
- Gumperz, John J. 1972. *Introduction*. In: Gumperz & Hymes (eds.), 1 - 25
- Gumperz, John J. 1974. *Linguistic anthropology in society*. *American Anthropologist* 76, 785 - 798
- Gumperz, John J. 1977. *Sociocultural knowledge in conversational inference*. In 28th Annual Round Table Monograph Series on Language and Linguistics, ed. by M. Saviile-Troike. Washington DC: Georgetown University Press
- Gumperz, John J. 1982 a. *Discourse strategies*. Cambridge: Cambridge University Press.
- Gumperz, John J. (ed.) 1982b. *Language and Social Identity*. Cambridge: Cambridge University Press
- Gumperz, John J. & D. Hymes (eds.). 1972. *Directions in Sociolinguistics. The Ethnography of Communication*. New York: Rinehart and Winston
- Gumperz, John J., G. Aulakh & H. Kaltmann. 1982. *Thematic structure and progression in discourse*. In Gumperz (ed.), 22 - 56
- Güntherodt, Ingrid. 1969. *Der Melodieverlauf bei Fragesätzen in zwei Lothringer Mundarten*. *Phonetica* 19, 156 - 169
- Güntherodt, Ingrid. 1971. *Der Tonhöhenverlauf bei Fragesätzen in Mundarten der Ostpfalz*. *Zeitschrift für Dialektologie und Linguistik* 38, 272 - 295
- Güntherodt, Ingrid. 1973. *A Prosodic Isogloss in German Dialects*. *Zeitschrift für Dialektologie und Linguistik* 1, 29 - 35
- Günthner, Susanne. 1993. *Diskursstrategien in der Interkulturellen Kommunikation*. Tübingen: Niemeyer
- Guy, Gregory, B. Horvath, J. Vonviller, E. Daisley and I. Rogers. 1986. *An intonational change in progress in Australian English*. *Language in Society* 15, 23 -52
- Hadding-Koch, Kerstin. 1956. *Recent works on intonation*. *Studia Linguistica* 10, 77 - 96
- Hall, Edwin T. 1955. *The anthropology of manners*. *Scientific American*. April, 84 - 90

- Halliday, Michael A.K. 1967a. Intonation and Grammar in British English. The Hague: Mouton
- Halliday, Michael A.K. 1967b. Notes on transitivity and theme in English, Part 2. *Journal of Linguistics* 3, 199 - 244
- Halliday, Michael A.K. 1970. A course in Spoken English: Intonation. London: Oxford University Press
- Halliday, Michael A.K. & Ruqaiya Hasan. 1976. Cohesion in English. London: Longman
- Hammond, Robert & Mevyn C. Resnick (eds.). 1988. Studies in Caribbean Spanish Dialectology. Washington, DC: Georgetown University Press
- Han, Yang Saxena. 1991. An Analysis of Particles in Mandarin. Tokyo: Simpusha
- Han, Yang Saxena. 1995. A pragmatic analysis of the *ba* particle in Mandarin Chinese. *Journal of Chinese Linguistics* 23 (2), 99 - 127
- Hardcastle, W. J. 1976. Physiology of Speech Production. London: Academic Press
- Henne, H. & H. Rehbein. ³1995. Einführung in die Gesprächsanalyse. Berlin: de Gruyter
- Henne, H. & O. Rongen et al. 1977. A handbook on Chinese Language Structure. Oslo: Universitetsforlag
- Henton, C. G. & A. W. Bladon. 1985. Breathiness in normal female speech: inefficiency versus desirability. *Language and Communication* 5, 221 - 227
- Hinnenkamp, V. 1989. Interaktionale Soziolinguistik und Interkulturelle Kommunikation. Tübingen: Niemeyer
- Hinnenkamp, V. 1997. Mißverständnisse in Gesprächen. Wiesbaden: Westdeutscher Verlag
- Hinton, L. , J. Nichols & J. Ohala (eds.). 1994. Sound symbolism. Cambridge: Cambridge University Press
- Hirano, M. & Kakita Y. 1985. Cover-body theory of vocal fold vibration. In Daniloff (ed.), 1 - 46
- Hirst, Daniel & Di Cristo (eds.). 1998. Intonation Systems. A survey of Twenty Languages. Cambridge: Cambridge University Press
- Hirst, Daniel & Albert Di Cristo. 1998a. A survey of intonation systems in: Hirst & Di Cristo (eds.), 1 - 44
- Ho, Aichen. 1977. Intonation variations in a Mandarin sentence for three expressions: Interrogative, exclamatory and declarative. *Phonetica* 34: 446 - 456
- Holmes, Janet. 1992. An Introduction to Sociolinguistics. London: Longman
- Holmes, Janet. 1995. Women, Men and Politeness. London: Longman
- Honikman, B. 1964. Articulatory settings. In Abercrombie et al. (ed.), 73 - 84
- Hong, B. 1985. Politeness in Chinese: Impersonal Pronouns and Personal Greetings. *Anthropological Linguistics* 2, 204 -213
- Horvath, Barbara. 1985. Variation in Australian English: A sociolinguistic study of English in Sydney. Cambridge: Cambridge University Press.
- Hu, Hsien-Chin. 1944. The Chinese concept of face. *American Anthropologist*, 1945 n.s. 46, 45 - 64
- Hymes, Dell. 1962. The ethnography of speaking. In: T. Gladwin & W. C. Sturtevant (eds.)
- Hymes, D. 1972a. Models of the Interaction of Language and Social Life. In Gumperz & Hymes (eds.), 35 - 71
- Hymes 1972b. Toward ethnographies of communication: the analysis of communicative events. In Giglioli (ed.), 21 - 44
- Hymes, D. 1974. Ways of speaking. Explorations in the ethnography of speaking. In Bauman and Sherzer (eds.), 433 - 451
- Isacenko, A. & Hans-Joachim Schädlich. 1970. A Model of Standard German Intonation. The Hague: Mouton
- Izard, C. E. (ed.) 1979. Emotions in personality and psychopathology. New York: Plenum

- Jackendoff, R. S. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Ma.: MIT Press
- Jarmen, E. & A. Cruttenden. 1976. Belfast intonation and the myth of the fall. *Journal of the International Phonetic Association* 6, 4 - 12
- Jefferson, Gail. 1972. Side Sequences. In Sudnov (ed.), 294 - 248
- Jefferson, Gail. 1973. A case of precision timing in ordinary conversation: overlapped tag-positioned address terms in closing sequences. *Semiotica* 9, 47 - 96
- Jefferson, Gail. 1978. Sequential aspects of storytelling in conversation. In Schenkein (ed.), 219 - 248
- Jefferson, Gail. 1983. Notes on some orderliness of overlap onsets. In d'Urso & Leonardi (eds.), 11 - 38
- Jefferson, Gail. 1984. On the organization of laughter in talk about troubles. In Atkinson & Heritage (eds.), 346 - 369
- Johnstone, Tom & Klaus R. Scherer. 1999. The effects of emotions on voice quality. In *Proceedings of the International Conference of Phonetic Sciences, 1999*. San Francisco, 2029 - 2032
- Jones, W.E. & J. Laver (eds.) 1973. *Phonetics in Linguistics. A book of readings*. London: Longman
- Karlgren, B. 1918. *A Mandarin phonetic reader in the Pekinese dialect*. Stockholm [Repr. 1975. Berlin: Springer]
- Kaufer, D. S. 1981. Understanding ironic conversation. *Journal of Pragmatics* 5 (6), 495 - 510
- Kenworthy, Joanne. 1978. The intonation of questions in one variety of Scottish English. *Lingua* 44, 267 - 282
- Kingdon, Roger. 1958. *The Groundwork of English Intonation*. London: Longman
- Klasmeyer, Gudrum & Walter F. Sendlmeier. 1995. Objective voice parameters to characterize the emotional content in speech. *Proceedings of the International Conference of Phonetic Sciences, 1995*. Stockholm. Vol. 1, 182 - 185
- Klatt, Dennis. 1975. Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics* 3, 129 - 140
- Klein, W. 1983. (ed.) *Intonation*. *Zeitschrift für Literaturwissenschaft und Linguistik*, LILI 49
- Kochman, Thomas. 1981. *Black and white styles in conflict*. Chicago: University of Chicago Press
- Kohler, Klaus. 1995. *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmitt Verlag
- Kotthoff, Helga & Ruth Wodak. 1997. *Communicating Gender in Context*. Amsterdam: John Benjamins
- Kramer, E. 1963. Judgement of personal characteristics and emotions from non-verbal properties of speech. *Psychological Bulletin* 60, 408 - 420
- Kratochvil, Paul. 1968. *The Chinese Language Today. Features of an Emerging Standard*. London: Hutchinson University Library
- Kratochvil, Paul. 1987. The case of the Third Tone. *Wang Li memorial volumes, English volume*, 253 - 276. Hongkong
- Kratochvil, Paul, 1998. Intonation in Mandarin Chinese. In: Hirst & Di Cristo (eds.), 417 - 431
- Kufner, Herbert L. 1962. *The Grammatical Structures of English and German*. Chicago: Chicago University Press
- Kühnen, S. 1977. *Stimme und Persönlichkeit: Analyse des Zusammenhangs von Stimmqualitätsvariablen und Persönlichkeitsvariablen*. Diploma Thesis, University of Giessen.
- Kvavik, K. 1988. Is there a Spanish imperative intonation? In Hammond & Resnick (eds.)
- Labov, William. 1966. *The social stratification of English in New York City*. Washington DC: Center for Applied Linguistics.

- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press
- Ladd, D. Robert. 1980. *The structure of intonational meaning. Evidence from English*. Bloomington: Indiana University Press
- Ladd, D. R. 1984. Declination: a review and some hypotheses. *Phonology Yearbook* 1, 53 -74
- Ladd, D. R., K.E.A. Silverman, F. Tolkmitt, G. Bergmann & K. R. Scherer. 1985. Evidence for the independent function of intonation contour type, voice quality and *F0* range in signaling speaker affect. *Journal of the Acoustical Society of America* 78 (2), 435 - 444
- Ladefoget, Peter. 1962. *Elements of acoustic phonetics*. Chicago: University of Chicago Press
- Ladefoget, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press
- Lakoff, Robin. 1975. *Language and Woman's Place*. New York: Harper and Row
- Lakoff, Robin. 1979. Stylistic Strategies Within a Grammar of Style. In Orasanu et al. (eds.), 53 - 78
- Lass, N. J. (ed.) 1979. *Speech and Language: advances in basic research and practice*. Vol. 1 & 2. New York: Academic Press
- Laver, John. 1968. Voice quality and indexical information. *British Journal of Disorders of Communication* 3, 43 - 54
- Laver, John. 1975. Individual features in voice quality. PhD Thesis, University of Edinburgh
- Laver, John. 1980. *The phonetic description of voice quality*. Cambridge: Cambridge University Press
- Laver, John & R. Hanson. 1981. Describing the Normal Voice. In Darby, J. K. (ed.)
- Leet-Pellegrini, H. M. 1980. Conversational dominance as a function of gender and expertise. In Giles et al. (eds.) 1980a, 97 - 104
- Lehiste, Ilse. 1975. The phonetic structure of paragraphs. In Cohen, A. & S. G. Nooteboom (eds.), 195 - 203
- Lehiste, Ilse. 1979. Perception of sentence and paragraph boundaries. In Lindblom, B. & S. Ohman (eds.)
- Leinonen, L., T. Hiltunen, I. Linnankoski & M.-L. Laakso. 1997. Expression of emotional-motivational connotations with a one-word utterance. *Journal of the Acoustical Society of America* 102, 1853 - 1863
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press
- Li, Charles, N. (ed.) 1976. *Subject and Topic*. New York: Academic Press
- Li, Charles, N. & Thompson, Sandra. 1981. *Mandarin Chinese. A Functional Grammar*. Berkeley: University of California Press
- Li, Aijun. 2002. Chinese Prosody and the Labelling of Spontaneous Speech. In *Proceedings of Speech Prosody 2002*, Aix-en-Provence, France, 39 - 44
- Lieberman, Philip. 1967. *Intonation, Perception and Language*. Research Monographs 38, MIT Press, Cambridge
- Lieberman, Philip. 1980. The innate, Central Aspect of Intonation. In Waugh & van Schooneveld (eds.), 187 - 200
- Lindblom, B. & S. Ohman (eds.). 1979. *Frontiers of Speech Research*. London: Academic Press
- Loman, B. 1975. Prosodic patterns in a Negro American dialect. In Enkvist & Ringbom et al. (eds.), 219 - 242
- Luchsinger, R. & G. E. Arnold. 1965. *Voice - speech - language. Clinical communicology: its physiology and pathology*. London: Constable
- Lyons, John. 1977. *Semantics*. (2 volumes) London & New York: Cambridge University Press

- Mallory, E. & V. Miller. 1958. A possible basis for the association of voice characteristics and personality traits. *Speech Monographs* 25, 255 - 260
- Malmberg, B. (ed.) 1968. *Manual of Phonetics*. Amsterdam: North Holland
- Maltz, Daniel & Ruth Borker. 1982. A cultural approach to male-female miscommunication. In Gumperz (ed.), 195 - 216
- Meyer-Eppler, Werner. 1957. Realization of Prosodic features in Whispered Speech. *Journal of the Acoustical Society of America* 29 (1), 104 - 106. Reprinted in Bolinger (ed.) 1972a
- Mair, Ch. & M. Markus (eds.) 1992. New departures in contrastive linguistics: Proceedings of the conference held in Innsbruck, Austria, 10-12 May 1991. (Two vols.) Innsbruck
- Markus, M. 1992. Rhythm, stress and intonation in English and German seen contrastively. In Mair & Markus (eds.), vol. 1, 21 - 36
- McConnell-Ginet, Sally. 1978. Intonation in the social context: Language and sex. Paper presented at the 9th International Congress of Sociology, Uppsala, Sweden
- McConnel-Ginet, S. 1983. Intonation in a Man's World. In Thorne & Henley (eds.), 69 - 87
- Mead, Richard. 1993. *Cross-cultural management communication*. Chichester: Wiley
- Milroy, Leslie. 1980. *Language and Social Networks*. Oxford: Blackwell
- Minifie, F. D., T. J. Hixon & F. Williams (eds.) 1973. *Normal aspects of speech, hearing and language*. Englewood Cliffs, NJ: Prentice-Hall
- Moore, G. E. 1939. Personality traits and voice quality deficiencies. *Journal of Speech Disorders* 4, 33 - 36
- Moosmüller, Silvia. 1997. The relevance of Fo for interruptions: A case study of political discussions in Austria. In Koffhoff & Wodak (eds.), 401- 421
- Morton, E. W. 1977. On the occurrence and significance of motivation- structural rules rules in some bird and mammal sounds. *American Naturalist* 111, 855 - 869
- Moulton, William G. 1962. *The Sounds of English and German*. Chicago: Chicago University Press
- Mozziconazzi, Sylvie. 1998. *Speech variability and emotion. Production and perception*. PhD. Diss. University of Eindhoven
- Mueller-Liu, P. 1997. *Family and Kinship in China and the West. A comparative study of German and Chinese Kinship Terms*. MS. University of the Saarland
- Murray, D. C. 1971. Talk, silence and anxiety. *Psychological Bulletin* 75, 244 - 260
- Murray, I. R. & J. L. Arnott. 1993. Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*, 93 (2), 1097 - 1108
- Neppert, J. & M. Pétursson. 1992. *Elemente einer akustischen Phonetik*. Hamburg: Helmut Buske
- Newmeyer, F.J. (ed.). 1989. *Linguistics. The Cambridge Survey. Vol. 4: Language: the Socio-Cultural Context*. Cambridge: Cambridge University Press
- Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press
- Norrick, Neal. 1991. On the organization of corrective exchanges in conversation. *Journal of Pragmatics* 16, 59 - 83
- Norrick, Neal. 1993. *Conversational Joking*. Bloomington: Indiana University Press
- Ochs, Elinor. 1973. A sliding sense of obligatoriness: the polystructure of Malagasy oratory. *Language in Society* 2, 225 -243
- Ochs Keenan, Elinor. 1974. Norm-makers, norm-breakers: use of speech by men and women in a Malagasy community. In Bauman & Sherzer (eds.), 125 - 143
- Ochs, Elinor, Emanuel Schegloff & Sandra M. Thompson (eds.) *Interaction and Grammar*. Cambridge: Cambridge University Press
- O'Connor, J. D. and G. F. Arnold. 1973. *Intonation of Colloquial English*. London: Longman.

- Öster, A. & A. Risberg. 1986. The Identification of the Mood of a Speaker by hearing-impaired Listeners. *Speech Transmission Laboratory - Q. Prog. Stat. Reports* 4, 79 - 90
- Ohala, John J. 1978. The production of tone. In Fromkin (ed.), 5 - 39
- Ohala, John J. 1983. Cross-language use of pitch: an ethological view. *Phonetica* 40, 1 - 18
- Ohala, John J. 1984. An ethological perspective on common cross-language utilization of F0 of voice. *Phonetica* 41, 1 - 16
- Ohala, John J. 1994. The frequency code underlies the sound-symbolic use of voice pitch. In Hinton, L., J. Nichols & J. Ohala (eds.), 325 - 347
- Ohala, John J. & M. Hirano. 1967. An experimental investigation of pitch change in speech. *Journal of the Acoustical Society of America* 42, 1208 - 1209
- Orasanu, J., M. K. Slater & L. L. Adler (eds.). 1979. *Language, Sex and Gender*. Annals of the New York Academy of Science, Vol. 327
- Osgood, C., G. Suci & P. Tannenbaum. 1957. *The measurement of Meaning*. Urbana, Ill.: University of Illinois Press
- Palmer, Jerry. 1994. *Taking Humour seriously*. London: Routledge
- Palmer, Harold. 1922. *English Intonation. With systematic exercises*. Cambridge: Heffner & Sons.
- Pellowe, J. & V. Jones. 1978. On the intonational variability of Tyneside speech. In Trudgill (ed.), 101 - 121
- Peters, J. 2001. *Frageintonation in der Pfalz. Eine Reanalyse des Güntherodt - Korpus*. MS. Universität Potsdam
- Pheby, John. 1975. *Intonation und Grammatik im Deutschen*. Berlin: Akademie Verlag
- Pheby, John. ²1984. *Phonologie: Intonation*. In Heidolph, K.E., W. Flämig and W. Motsch. *Grundzüge einer deutschen Grammatik*. Berlin: Akademie Verlag, 839 - 897
- Pierrehumbert, J. B. 1980. *The Phonology and Phonetics of English Intonation*. PhD diss., MIT
- Pike, Kenneth. 1945. *The Intonation of American English*. University of Michigan Publications, Linguistics 1, Ann Arbor, Mi.
- Pike, Kenneth. 1947. *Language in Relation to a Unified Theory of Behavior*. The Hague: Mouton.
- Pike, Kenneth. 1948. *Tone Languages*. University of Michigan Publications, Linguistics 4, Ann Arbor, Mi.
- Psathas, G. (ed.) 1979. *Everyday Language. Studies in ethnomethodology*. New York: Free press
- Protopapas, A. & P. Lieberman. 1997. Fundamental frequency of phonation and perceived emotional stress. *Journal of the Acoustical Society of America* 101 (4), 2267 - 2277
- Pützer, Manfred & W. J. Barry. 2003. *The Usability or Perceptual Ratings of Voice Quality*. Paper given at the Symposium for the 60th birthday of Prof. W.J. Barry. April 25, 2003
- Quirk, R., A.P. Duckworth, J. Svartvik, J.P.L. Rusiecki & A.J.T. Colin. 1964. Studies in the Correspondance of prosodic to grammatical features of English. *Proc.Ling.* IX: 679 - 691
- Reichardt, Manfred & Shuxin Reichardt. 1990. *Grammatik des modernen Chinesisch*. Leipzig: Verlag Enzyklopädie
- Reisman, Karl. 1974. Contrapuntal conversations in an Antiguan village. In Bauman & Sherzer (eds.), 110 - 124
- Riesel, E. 1963. *Stilistik der deutschen Sprache*. Moskow
- Robins, R. H. ²1979. *A Short History of Linguistics*. London & New York: Longman
- Rochester, S. R. 1973. The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research* 2, 51 - 81
- Romaine, Suzanne. 1978. Postvocalic /r/ in Scottish English: a sound change in progress? In Trudgill (ed.), 144 - 157

- Romaine, Suzanne (ed.) 1982. *Sociolinguistic Variation in Speech Communities*. London: Edward Arnold
- Rumjancev, M. K. 1972. *Tone and Intonation in Modern Chinese*. Moskow: Izdatel'stvo Moskovskogo Universiteta (Reviewed by A. V. Lyonin in *Journal of Chinese Linguistics* 6 (1978), 120 -168)
- Ross, Alison. 1998. *The Language of Humour*. London: Routledge
- Russ, Ch. V. J. (ed.) 1981. *Contrastive aspects of English and German*. Heidelberg: Groos
- de Saussure, Ferdinand. 1916/1962. *Cours de Linguistic Générale*. Paris: Payot
- Saville-Troike, Muriel, 1982. *The Ethnography of Communication. An Introduction*. Oxford: Basil Blackwell
- Sacks, Harvey, 1984. Notes on methodology. In Atkinson & Heritage (eds.), 21 - 27
- Sacks, H., E. Schegloff & G. Jefferson. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50, 696 - 735
- Schädlich, H.-J. & H. Eras. 1969. Vergleichende Untersuchungen über die Satzintonation in deutschen Dialekten. In: *Zur phonetischen und phonologischen Untersuchung prosodischer Merkmale*. Berlin: Deutsche Akademie der Wissenschaften, 27 - 46
- Schegloff, E. & H. Sacks. 1973. Opening up Closings. *Semiotica* 8, 289 - 327
- Schenkein, J. (ed.). 1978. *Studies in the Organization of Conversational Interaction*. New York: Academic Press
- Scherer, Klaus R. 1970. Attribution of Personality from voice: a cross-cultural study on the dynamics of interpersonal perception. PhD. Thesis. Harvard University
- Scherer, Klaus, R. 1974. Voice quality analysis of American and German speakers. *Journal of Psycholinguistic Research* 3, 281 - 190
- Scherer, Klaus, R. 1979a. Nonlinguistic vocal indicators of emotion and psychopathology. In: Izard, C. E. (ed.), 495 -529
- Scherer, Klaus R. 1979b. Personality markers in speech. Scherer, Klaus R. & Giles, Howard (eds.), 147 - 209
- Scherer, Klaus R. 1981. Speech and emotional states. In: Darby, J. K. (ed.), 189 - 220
- Scherer, Klaus R. 1986. Vocal affect expression: A Review and a Model for Future Research. *Psychological Bulletin* 99, 143 - 165
- Scherer, Klaus R. 1995. Expression of Emotion in Voice and Music. *Journal of Voice*, 9 (3), 235 - 248
- Scherer, Klaus, H. London & J. J. Wolf. 1973. The voice of confidence: Paralinguistic cues and audience evaluation . *Journal of Research in Personality* 7, 31 - 44
- Scherer, Klaus R. & Howard Giles (eds.) 1979. *Social markers in speech*. Cambridge: Cambridge University Press
- Scherer, K. R., D. R. Ladd & K.E.A. Silverman. 1984. Vocal cues to speaker affect: testing two models. *Journal of the Acoustical Society of America*. 76, 1346 - 1356
- Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge and New York: Cambridge University Press
- Schiffrin, Deborah. 1989. Conversational Analysis. In Newmeyer (ed.), 251 - 276
- Schmidt, S. 1974. (ed.) *Pragmatik II*. München: Fink
- Schröder, P. & H. Steger. (eds.) 1981. *Dialogforschung*. Düsseldorf: Pädagogischer Verlag Schwann
- Schubiger, Maria. 1958. *English Intonation. Its form and function*. Tübingen: Niemeyer
- Scollon, Ron. 1985. The machine stops: Silence in the metaphor of malfunction. In Tannen & Saville-Troike (eds.), 21 - 30
- Scollon, Ron. 1995. *Intercultural Communication. A Discourse Approach*. Oxford: Blackwell
- Searle, John. 1969. *Speech Acts. An Essay in the philosophy of language*. Cambridge: Cambridge University Press

- Searle, John. 1975. Indirect Speech Acts. In Cole, P. & J. L. Morgan (eds.), 59 - 82
- Selting, Margret, 1995. Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation. Tübingen: Niemeyer
- Shangwu Yinshuguan. Xihua Zidian. 1989. Hongkong. Commercial Press
- Shen, Susan, Xiao-Nan. 1990. The Prosody of Mandarin Chinese. Berkeley: University of California Press
- Sherzer, Joel. 1974. Namakke, Summakke, Kormakke: three types of Cuna speech event. In Bauman & Sherzer (eds.), 263 - 282
- Siegmán, A. W. 1978. The tell-tale voice: nonverbal messages of verbal communication. In Siegmán, A. W. & S. Feldstein (eds.)
- Siegmán, A. W. & S. Feldstein (eds.) 1978. Nonverbal Behavior and Communication. Hillsdale, N.J.
- Siegmán, A. W. & B. Pope. 1965. Effects of question specificity and anxiety-producing messages on verbal fluency in the initial interview. *Journal of Personality and Social Psychology* 2, 255 - 230
- Sievers, E. 1893. Grundzüge der Phonetik. Leipzig
- Smith, P. & Z.-M. Wang. 1996. Chinese Leadership and Organizational Structures. In Bond (ed.), 322 - 337
- Smith, Philip. 1985. Language, the Sexes and Society. Oxford: Basil Blackwell
- Stubbs, Michael. 1983. Discourse Analysis: The Sociolinguistic Analysis of Natural Language. Chicago: University of Chicago Press
- Sudnov, D. (ed.). 1972. Studies in Social Interaction. New York: Free Press
- Sundberg, J. (ed.) 1981. Research Aspects on Singing, 33. Stockholm: Royal Swedish Academy of Music
- Szemerényi, Oswald. 1989. Einführung in die Vergleichende Sprachwissenschaft. Darmstadt: Wissenschaftliche Buchgesellschaft
- Tannen, Deborah, (ed.) 1982. Analyzing Discourse: Text and Talk. Georgetown University Round Table on Language and Linguistics, 1981
- Tannen, Deborah. 1981. New York Jewish Conversational Style. *International Journal of the Sociology of Language* 30, 133 - 139
- Tannen, Deborah. 1984. Conversational Style: Analyzing Talk Among Friends. Norwood, NJ: Ablex
- Tannen, Deborah. 1986. That's not what I meant. How Conversational Style makes and breaks your Relations with Others. New York: Morrow
- Tannen, Deborah. 1989. Talking Voices. Repetition, dialogue and imagery in conversational discourse. Cambridge and New York: Cambridge University Press
- Tannen, Deborah. 1990. You just don't understand. Women and Men in Conversation. New York: Morrow
- Tannen, Deborah (ed.) 1993a. Framing in Discourse. New York and Oxford: Oxford University Press.
- Tannen, Deborah. 1993b. Introduction. In Tannen (ed.). 1993c
- Tannen, Deborah. (ed.) 1993c. Gender and Conversational Interaction. New York and Oxford: Oxford University Press
- Tannen, Deborah. 1994. Gender and Discourse. New York and Oxford: Oxford University Press
- Tannen, Deborah & M. Saville-Troike (eds.). 1985. Perspectives on Silence. Norwood, NJ: Ablex
- Tao, Hong-Yin. 1996. Units in Mandarin Conversation. Prosody, Discourse and Grammar. Amsterdam and Philadelphia: John Benjamins
- Tarone, E. S. 1976. Aspects of intonation in Black English. *American Speech* 48, 29 - 36

- Tartter, V. C. 1980. Happy Talk: Perceptual and Acoustic Effects of Smiling on Speech. *Perceptory Psychophysics*. 27, 24 - 27
- Tench, Paul. 1990. *The Roles of Intonation in English Discourse*. Frankfurt: Lang
- Tench, Paul. 1995. *The Intonation Systems of English*. London: Cassell
- Terango, L. 1966. Pitch and duration characteristics of the oral reading of males on a masculinity-femininity dimension. *Journal of Speech and Hearing Research* 9, 590 - 595
- Thorne, Barrie & Nancy Henley (eds.) 1975. *Language and Sex: Difference and Dominance*. Rowley: Newbury House
- Thorne, Barrie, Cheris Kramarae & Nancy Henley (eds.) 1983. *Language, Gender and Society*. Rowley: Newbury House
- Thompson, J. B. 1991a. Introduction. In Thompson 1991b, 1 - 31
- Thompson, J. B. (ed.) 1991b. *Pierre Bourdieu: Language and Symbolic Power*. Cambridge Ma.: Harvard University Press
- Thorsen, Nina. 1980. A study of the perception of sentence intonation: Evidence from Danish. *Journal of the Acoustical Society of America* 67, 1014 - 1030
- Trim, J.L.M. 1959. Major and Minor tone-groups in English. *Le maitre phonétique*, 3rd series, No. 112, 26 - 29. Also in Jones & Laver (eds.) 1973
- Trimboli, F. 1973. Changes in voice characteristics as a function of trait and state personality variables. *Dissertation Abstracts International* 33, 3965
- Trojan, F. 1952. *Der Ausdruck der Sprechstimme*. Wien-Düsseldorf
- Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1, 179 - 196
- Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. London: Cambridge University press
- Trudgill, Peter. 1978 (ed.) *Sociolinguistic Patterns in British English*. London: Edward Arnold
- Trubetzkoy, N. 1939. *Grundzüge der Phonologie*. *Travaux du cercle linguistique de Prague* 7, Prague [Repr. Göttingen 1958]
- Uldall, Elizabeth. 1960. Attitudinal meanings conveyed by intonation contours. *Language and Speech* 3, 223 - 234
- Uldall, Elizabeth. 1964. Dimensions of meaning in intonation. In Abercrombie et al. (eds.), 271 - 279; Reprinted in Bolinger (ed.) 1972, 250 - 259
- Ullan, R. 1969. Some general characteristics of interrogative systems. *Working Papers in Language Universals*, Stanford. 1, 39 - 63a
- d'Urso, V. & P. Leonardi. (eds.) 1983 *Discourse Analysis and Natural Rhetorics*. Padova: Cleup Editore
- Vanderslice, R. & L. S. Pierson. 1972. Prosodic features of Hawaiian English. In Bolinger 1972a, 439 - 450
- Van Dyke, Teun. 1980. *Macrostructures: An Interdisciplinary Study of Global Structures in Discourse, Interaction and Cognition*. Hillsdale, NJ: Erlbaum
- Waugh, Linda R. & C. H. van Schooneveld (eds.). 1980. *The Melody of Language*. Baltimore: University Park Press
- Webster, N. 1789. *Dissertations on the English language*. Boston: Isaiah Thomas
- Williams, C. E. & K. N. Stevens. 1972. Emotions and Speech: Some Acoustical Correlates. *Journal of the Acoustical Society of America* 52, 1238 - 1250
- Yang, Martin C. 1945. *A Chinese Village*. New York: Columbia University Press
- Yang, Li-Chiung & N. Campbell. 2001. Linking Form to Meaning: The Expression and Recognition of Emotions through Prosody. In *Proceedings of the 4th Speech Synthesis Workshop*, Scotland, 2001. 141f.
- Yang, Yufang & Wang, Bei (2002). Acoustic Correlates of Hierarchical Prosodic Boundary in Mandarin Chinese. *Speech Prosody*. Aix-en-Provence. 707 - 710

- Yao, Qian & Pan, Wuyun. 2002. Prosodic Word: The Lowest Consituent in the Mandarin prosody Processing. In Proceedings of Speech Prosody 2002, Aix-en-Provence. 591 - 594
- Yates, J.F. & J.-W. Lee. 1996. Chinese Decision-making. In Bond (ed.), 338 - 351
- Young, Linda W. Y. (1982) Inscrutability revisited. In Gumperz (ed.) 1982b, 72 - 84
- Yngve, V. 1970. On getting a word in edgewise. Papers from the Sixth Regional meeting of the Chicago Linguistics Society. Chicago, 567 - 578
- Yule, George. 1980. Intonation and givenness in spoken discourse. *Studies in Language* 4, Amsterdam, 271 - 286
- Yuan, Jiahong, Chilin Shih & G. P. Kochanski. 2002. Comparison of Declarative and Interrogative Intonation in Chinese. In Proceedings of Speech prosody 2002, Aix-en-Provence. 711 - 714
- Zee, E. 1980. A spectrographic investigation of Mandarin tone sandhi. *UCLA working papers in phonetics* 49, 98 - 116
- Zemlin, W. R. 1964. *Speech and hearing science*. Champa, Il: Stipes

