TEMPO AND PROSODIC PATTERN IN CHINESE AND SWEDISH

Eva Gårding and Jialu Zhang

Abstract
This paper is a study of how speech rate influences temporal and tonal patterns in Standard Chinese. It includes a comparison with an earlier, similar study of Swedish. The material, sentences with tonal (accentual) patterns in focus-free statement intonation, was produced by several speakers under three rates of speech, labelled Normal, Fast and Slow. The main effect of a shift of tempo from slow to fast is a compression and an upward movement of the overall pitch range while the number of turning points and their positions relative to the segments are well retained. In the time domain there is a lengthening of about 50% from normal to slow and a shortening of about 25% from normal to fast speech. This compression is not uniform. For the higher tempo the last constituent increases in relative duration and prominence at the expense of the segments of the first constituent. The comparison with Swedish shows basic similarities. Observed differences can be explained by reference to the larger number of pitch-determined syllables in a Chinese sentence as compared to a Swedish one.

Introduction

The main purpose of this study is to show how speech rate affects prosodic patterns in Chinese and Swedish.

Speech rate as a variable was used by Kozhevnikov and Chistovich with a view to investigating the organisation of articulatory movements on a neurological level (1965). By measuring durations in Russian sentences between sharply defined articulatory events at different rates of speech, they found that the segments of certain sequences, e.g. the word and the syllable, have constant relative durations. They concluded that speech rate must be below the articulatory program of such sequences.

The view of the syllable as a complex articulatory gesture was also adopted by Gårding who studied the manifestations of various boundary signals in different styles and rates of speech (Gårding 1967 p.23 ff.).

Recently a group of researchers at Haskins Laboratories have used 'gestures', i.e. a limited set of specified coordinated articulatory movements, as primitive units in their development of a speech synthesis model (Browman et al. 1986). Within this framework dynamic properties in speech caused by shifts of tempo and style are being analysed.

The suprasegmental level has received much less attention. One previous study dealt with the effect of speech rate on temporal and tonal patterns in different Swedish dialects (Gårding 1975). With support from an electromyographic study (Gårding et al. 1975) the notion of laryngeal gestures and their timing was used to explain the effect of tempo on the pitch curve. Our present paper deals with Standard Chinese and a comparison wih Swedish makes it possible to sort out general tempo effects in the speech signal from phonological peculiarities caused by tones or accents. It also shows that the notion of laryngeal gestures and their overlap in fast speech is fruitful to explain tonal targets and tempo effects in Chinese as well as Swedish.

After a presentation of the material and the method which are similar for both analyses the main section will be devoted to our recent analysis of Standard Chinese. The results are compared with observations in the earlier Swedish investigation.

1. Material, method and terminology

For Swedish the material consisted of various noun phrases in which the noun and the preceding adjective had four combinations of accent 1 and accent 2 (see Gårding 1975). For Chinese the following two sentences were used:

(1) Sùng Yán mài niúròu 'Sung Yan sells beef' with alternating falling and rising tones

(2) Sùng Yàn mài mòli 'Sung Yan sells jasmin' with falling tones only, the last syllable being atonic.

Phonologically the falling tone is analysed as HL (high low), and the rising tone as LH (low high)(see Gårding et al. 1983). The abbreviations may be regarded as the targets of a particular tone.

Behind the tonal movements there is a laryngeal gesture, i.e. a coherent cycle of activity of some laryngeal muscles in particular the cricothyroid (for an EMG investigation of Swedish accents, see Gårding et al. 1975).

The Chinese informants represented Standard Chinese and the Swedish informants three well-defined dialects. The test sentences were read from cards in focus-free statement intonation at three rates of speech, labelled Normal, Fast and Slow. The instructions were: Start with normal. For Slow think of a situation where you have to dictate slowly without ruining the coherence of the sentence. For Fast imagine that we have very little tape left. The test sentences were read three times by the informants.

In the acoustic analysis, segment durations and fundamental frequency were measured from broad- and narrow-band spectrograms. The Chinese material was also analysed by speech wave, intensity and fundamental frequency records obtained from the digitized speech wave by means of the ILS program.

Special attention was paid to parameters and concepts that had proved useful in cross-linguistic analyses of intonation (Gårding 1981 and 1984): the turning points of the F0 curve (i.e. the local maxima and minima of the curve) with their positions relative to segments and frequency scale, the notion of turning point fixation (which means that certain turning points have rather fixed locations relative to segments regardless of spech act), and the tonal grid as an expression of sentence intonation. A tonal grid is a set of near-parallel lines that enclose normal-size tones or accents.

In our analysis we shall make a distinction between 'rhythmic' which refers to the subjective impression of the time relations of a pattern and is transcribed by using marks for short (∪) and long (-), and 'temporal' which refers to objective measurements.

The term sentence prominence denotes the phrase-level prominence of Chinese and sentence accent denotes the corresponding phenomenon in Swedish.

2. Chinese

Temporal pattern

Figures 1 and 2 show examples of the raw data obtained by the ILS program. With a shift of tempo there is an overall lengthening of 50% for slow speech and a shortening of 25% for the fast speech of both speakers. There is more variation in lengthening than in shortening.

In Figure 3 the segments of the utterances of the three tempi have been brought to comparable durations by normalized scaling. Since the lines connecting the segments are not vertical, expansion and compression of the normal-rate utterance have not been uniform. In other words, the time scale has been warped.

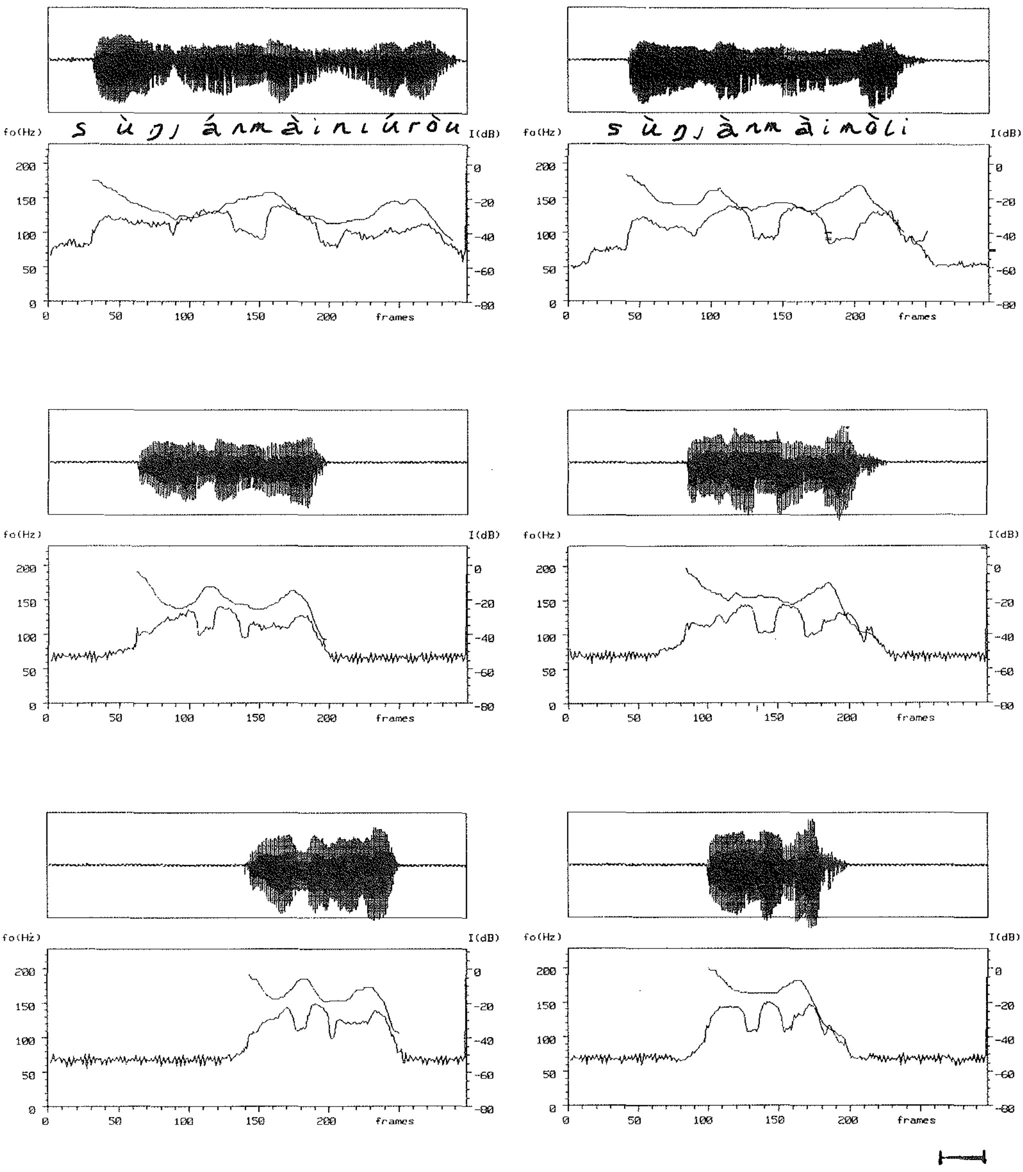


Fig.1. Waveform, FO and intensity tracings of Sùng Yán mài niúròu left and Sùng Yàn mài mòli right in three tempi, from top to bottom, slow, normal, fast. Speaker 1.

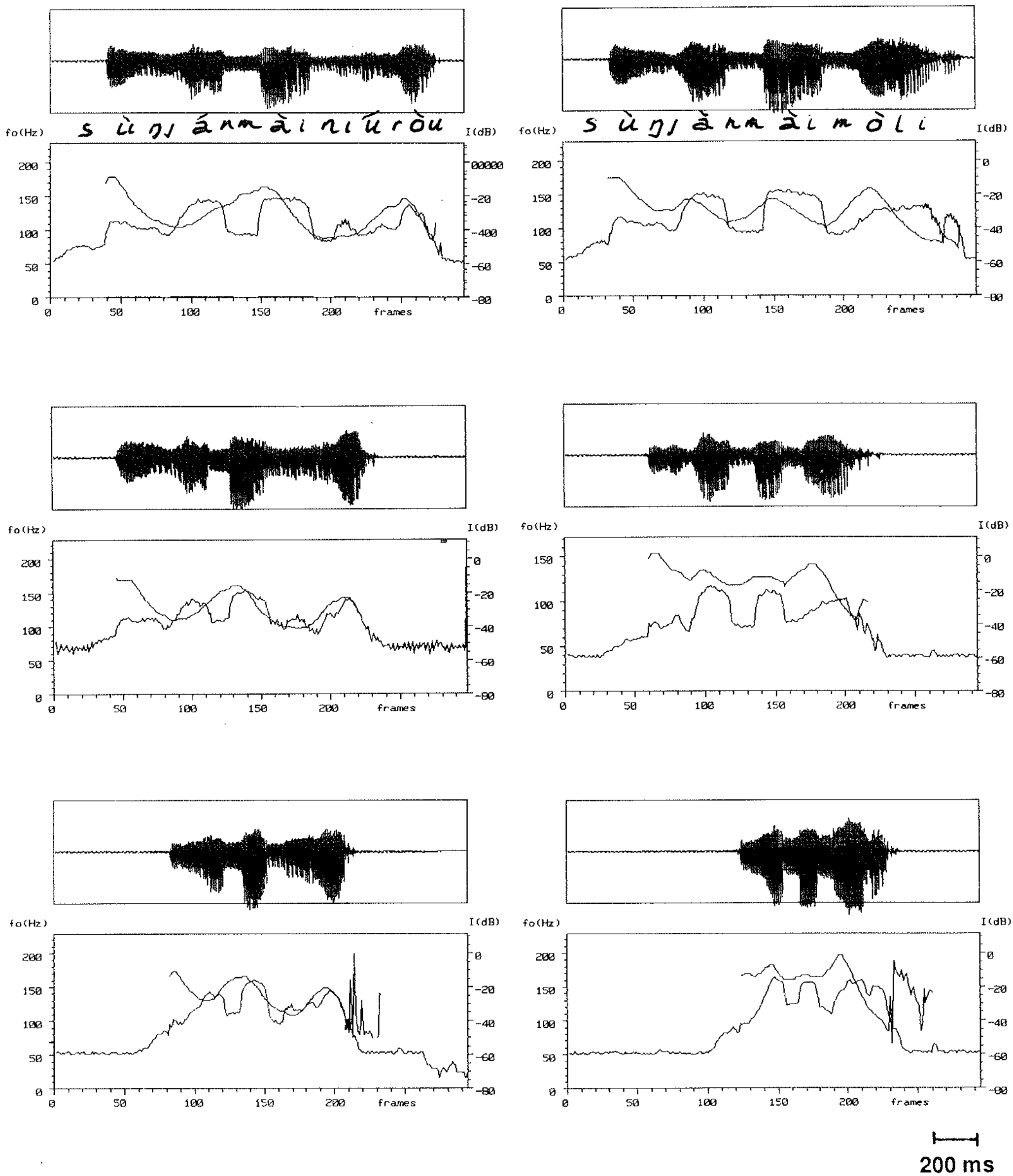


Fig.2. Waveform, F0 and intensity tracings of Sùng Yán mài niúròu left and Sùng Yàn mài mòli right in three tempi, from top to bottom slow, normal, fast. Speaker 2.

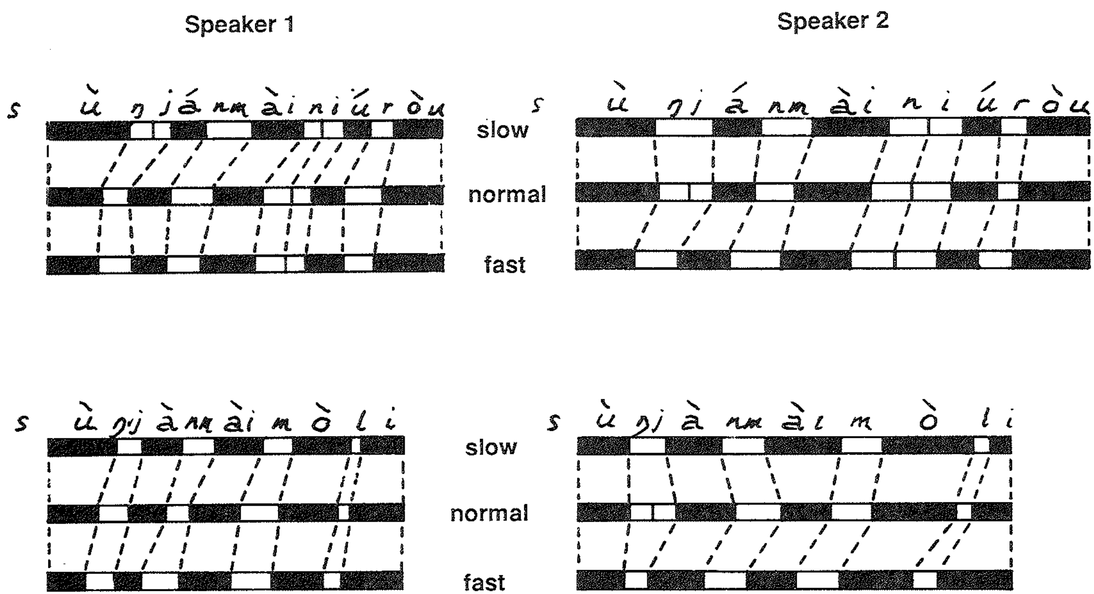


Fig.3. Relative durations of segments in three tempi, slow, normal fast. Sentence 1 is above and Sentence 2 below. Speaker 1 is to the left and Speaker 2 to the right.

With the arrangement chosen in the presentation most lines are tilted to the right, which means that with increased tempo the early part of the sentence gets more compressed than the latter part which has a corresponding expansion. In this process the first constituent of the phrase, the subject, goes down from about half of the utterance to about one third, and the last constituent, consisting of the compounds niúròu and mòli, changes correspondingly. In fast speech the compounds occupy half of the duration of the utterance. Mainly responsible for this change is probably the sentence prominence (comparable to the sentence accent in Swedish and other accent languages) which has a more important role in fast speech and is carried by the last tonal morpheme, i.e. by -ròu and mò-, according to the rules of this dialect. The predicate verb, mài, has a uniform part of the utterance, regardless of tempo.

For the first sentence the auditive rhythmic pattern may be represented as:

Sūng Yăn māi niŭrōu

and for the second

Sūng Yăn măi mōlĭ

The acoustic record suggests that the impression of prominence is a combined effect of relative duration, intensity and pitch range. There seem to be trading effects in neighbouring syllables belonging to the same phrase. The lengthening of mò- under sentence prominence may be the cause of the short duration of the preceding mài. (Compare the corresponding mài of the other sentence which is further away from the sentence prominence and longer.) The interaction of temporal and tonal patterns needs a special investigation.

Tonal pattern

In Figure 4 the normal-rate temporal pattern was used as a base and the tonal patterns of the three tempi were transposed to their respective segments via the turning points. The fast renderings are generally a bit higher and have a smaller range, at least in the falling-rising pattern.

In one case, Sùng Yàn mài mòli, Fast, the number of turning points has changed with tempo. The two falls of the name Sùng Yàn are exchanged for a single fall which makes it comparable to the compound mòli appearing in the same sentence. However, mòli has one fall regardless of tempo.

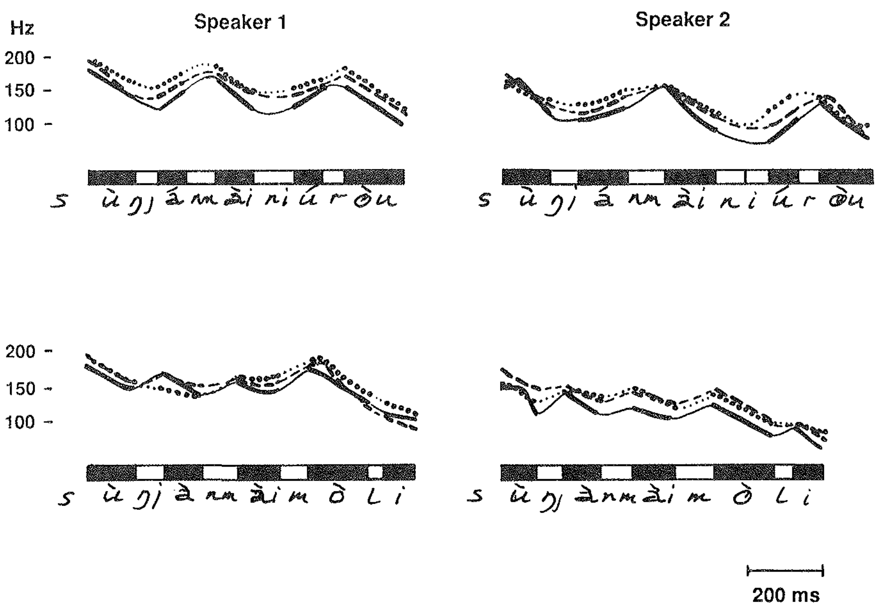


Fig.4. Tonal patterns of three tempi transposed to the normal-rate temporal pattern. Unbroken line is for slow performance, broken line for normal and dotted for fast. Sentence 1 is above and Sentence 2 below. Speaker 1 is to the left and Speaker 2 to the right.

In all other cases the number of turning points is retained and they have rather fixed locations relative to the segments. In most cases the first target of a tone is reached at the beginning of the vocalic segment in close connection with the C/V boundary. However, in fast speech there is a tendency for the first target point to be reached in the preceding consonant rather than at the beginning of the vocalic segment. See Figures 1 and 2.

There is covariation in the positions of the turning-points. This means that a contour is retained in shape even if it is shifted up or down in pitch.

In sentence 1 the falling and rising tones have about the same ranges for a particular tempo. However, in sentence 2, the sequence of falls, the middle tones have compressed ranges in accordance with a sandhi rule, stating that a falling tone becomes reduced to half size before a subsequent falling tone. This compression may be due to undershoot (see below), which can be explained by the physiological difficulty of producing two contour tones of he same kind after each other in a short time span. Since it is the vocalic segment that carries the perceptually pertinent tone pattern (Kratochvil 1968, Howie 1974, Gårding et al. 1987), the laryngeal muscles have to make a very quick manoeuvre at best in the consonant(s), in this case from the low target of one tone to the high target of the next. From this point of view, sentence 1 with its falling and rising tones gives the speaker a much easier tonal task since the second target of the first tone coincides with the first target of the next.

In Sùng Yàn of Sentence 2 (Speaker 1, Fast) there is no individual trace of the second tone and before sentence prominent mòli there is a levelling of the falling tone which may be caused by gesture overlap. Speaker 2 is not comparable in this respect since he changes the rhythmic pattern of his fast utterance.

Cut out from its context, this and similar fast utterances would be ambiguous. In real life, however, such utterances hardly exist in isolation. Coarticulation does not endanger communication in general.

Laryngeal pattern

The turning points are associated with changes in the activity pattern of the pitch-controlling muscles, in particular the cricothyroid muscles. Hence the distances between the high turning points, which have a more precise location than the low ones, indicate how tempo influences the train of laryngeal gestures. A list of these measurements (not given here) shows that the relative distance between the two peaks of <u>Sùng Yán</u> in Sentence 1 decreases with increased speed. In other words the gestures behind the semantically close-knit names come closer together than those controlling <u>mài niúròu</u> i.e. the predicate verb and the object. These results are of course just another aspect of the turning point fixation and the observed changes in the temporal pattern.

Sentence intonation (disregarding tempo)

Since the speakers have used the same declarative intonation in both sentences it is possible to abstract it from the curves.

The similarity of the six renderings of each sentence and the covariation permits us to exhibit them simultaneously in one diagram. In Figure 5, the left part of which refers to Sentence 1 produced by Speaker 1, the intonation curves are shown together with an average normal-rate temporal pattern as a base. All renderings lie within a band limited by the small vertical intervals. A curve has been interpolated through the means of each interval. We regard this curve as an expression of the target sentence intonation, i.e. the intonation that the speaker aimed at in all his productions of this sentence. This intonation is clearly falling and since the first and the last intervals are not broader than the others we can infer that the slope is determined by the length of the utterance. The combined effect of the sentence intonation and the tonal pattern can be given a concrete representation by a common tonal grid, which is illustrated by the broken lines in the figure.

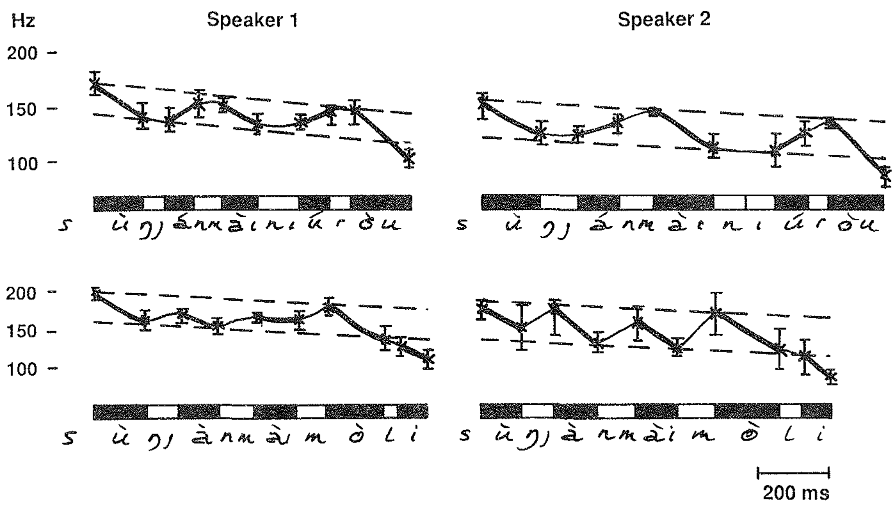


Fig.5. Intonation ranges and means of six renderings of each sentence imposed on a common normal-rate temporal pattern. All renderings lie within a band limited by the small vertical intervals. A curve has been interpolated through the means of each interval. Broken lines illustrate the tonal grid.

In the right-hand part of Figure 5, the renderings of sentence 2 have been treated in the same way and the grid of sentence 1 has been imposed. The sentence prominence is in both cases rendered by a large amplitude of the tonal movement. In the second sentence the sentence prominence falls on the penultimate syllable, the last one being atonic. In spite of the different locations of the sentence prominence the grid fits rather well, and the statement intonation used for both sentences can be expressed by the same falling grid. For sentence 2, due to the tone sandhi rule mentioned earlier, only the first and the last tones reach the topline of the grid.

A similar grid for the different renderings of the two sentences have been obtained for the other speaker. One difference is that his speech shows less coarticulation. Regardless of tempo his falling tones remain falling.

Grids adjusted to tempo

In the preceding section tempo was disregarded. When tempo is taken into account, the following small changes in the intonation curves are apparent. With increased rate of speech the entire curve is shifted up in frequency by a small amount and the range is compressed. The appendix is an attempt to simulate these modifications in a numerical generative scheme.

Summary of the Chinese analysis

Judging from this material the following prosodic features stay invariant through the three tempi.

(1) The sentence prominence.

(2) The tonal pattern.

(3) The sentence intonation.

Invariance is here taken in the phonological sense. There may be modifications in the signal but there is no loss of identity. The speaker obviously intends to retain the communicative value of a sentence even if he speaks fast.

The following phonetic modifications have been observed.

(1) Although the binary rhythmic division of the sentence into strong and weak is still present in fast speech, the temporal pattern changes so that more relative duration is given to the last constituent (niùróu, mòli) which enhances its perceptual importance and may motivate a third degree of prominence in a transcription.

(2) With a shift of tempo the position of the tonal pattern relative to the segments has a certain leeway. The first target point, represented by the first high or low turning point of the syllable is more fixed than the second one. It is closely connected with the C/V boundary but has a tendency to move leftward into the preceding consonant in fast speech. The position of the second target point is somewhat adjusted to meet the requirements of the following syllable and is therefore more variable.

There may be complete tone reduction in syntactically and semantically close knit syllables in fast speech.

(3) The overall range of sentence intonation gets smaller and moves up on the frequency scale.

The observed tempo effects are illustrated by the pitch curves of Figures 6 and 7, derived from our two speakers' slow and fast renderings of the test sentences. Below each observed curve is a schematized construction which follows the model for intonation described elsewhere (e.g. Gårding et al. 1984. The main principles are given in the appendix). The schematized curves have been imposed on the digitized originals by means of the ILS

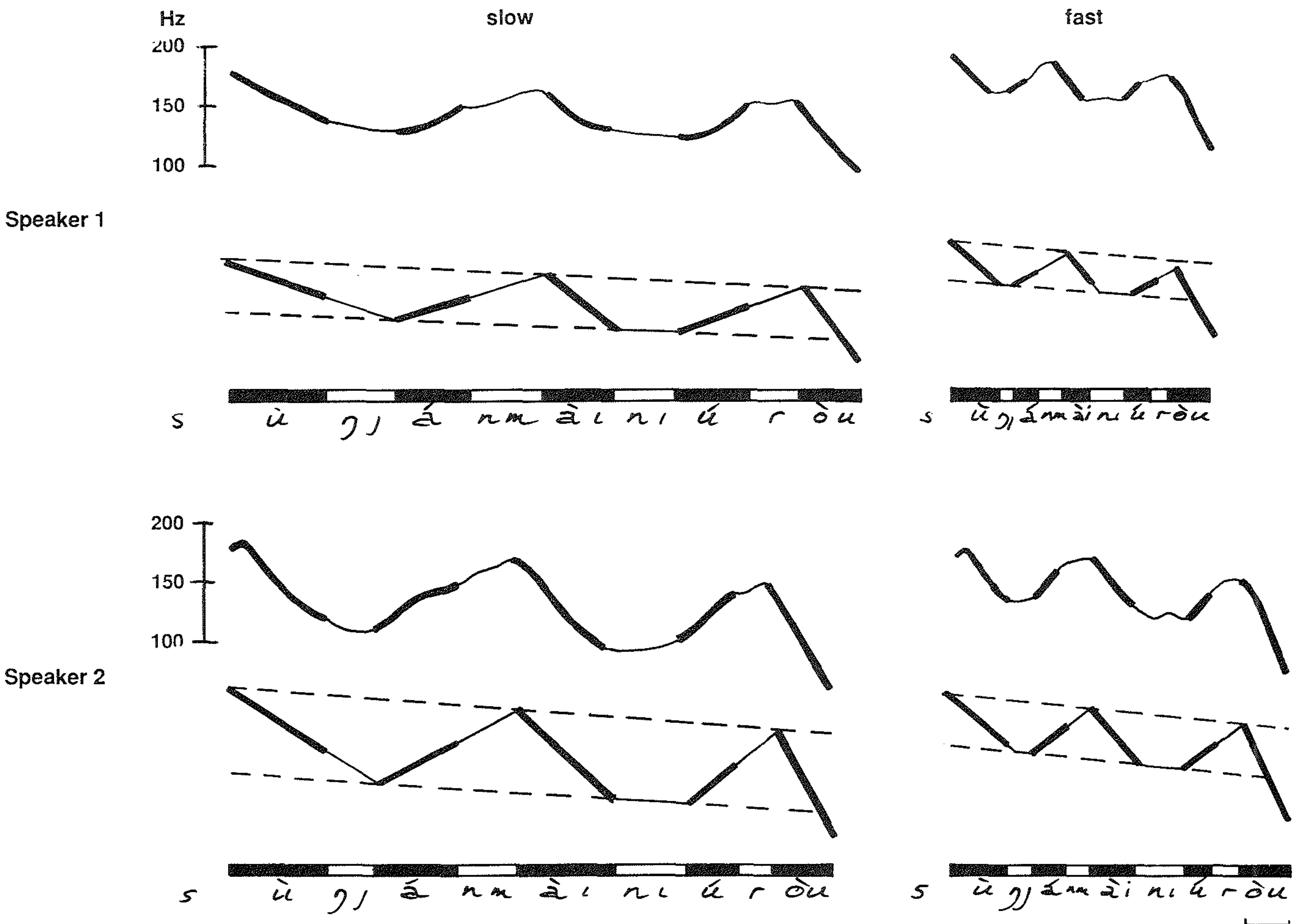


Fig.6. Observed and generated F0 curves of Sentence 1 in the slow and fast performance of Speaker 1 above and Speaker 2 below.

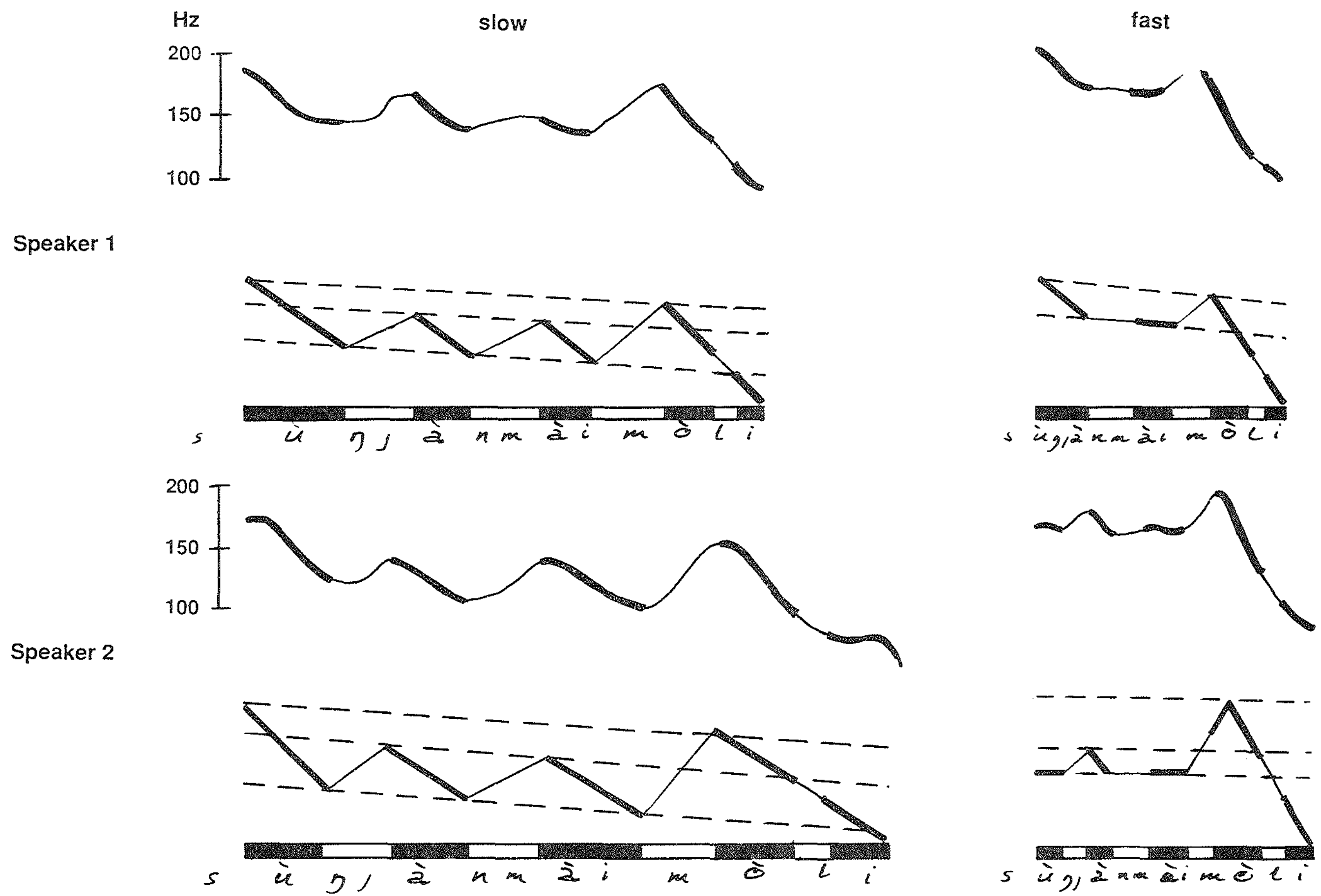


Fig.7. Observed and generated F0 curves of Sentence 2 in the slow and fast performance of Speaker 1 above and Speaker 2 below.

program. In spite of the straight-line interpolations and the other simplifications the resulting utterances are hardly distinguishable from the original ones. Another result of our ILS experiments is that the characteristics of fast speech, narrower and raised grids imposed on the normal-rate sentences, were not sufficient to create an impression of increased tempo. A change in the time domain seemed necessary.

3. Swedish

Only the most important findings in the earlier paper (Gårding 1975) will be summarized. Three dialects were considered, Southern Swedish (Helsingborg, Skåne), Central Swedish (Stockholm) and Eastern Swedish (Helsinki). Southern Swedish, being spoken on a substratum of Eastern Danish, is prosodically very different from the central dialects, e.g. the sentence accent is rising in the central dialects and falling in the south (Bruce & Gårding 1978). Lacking distinctive accents, Finland Swedish is different from both.

The sentence used for demonstration is <u>en manli(g) nunna</u> 'a masculine nun' with A2, transcribed by a grave accent mark, on both the noun and the adjective.

Figure 8 shows average durations and fundamental frequency curves of the demonstration sentence spoken at three rates of speech by representatives of three dialects.

Temporal pattern

Figure 9 shows the relative durations of the segments of the demonstration sentence in focus-free declarative intonation, spoken by three informants across three tempi. We can infer from the Central Swedish renderings that the last syllable, in this

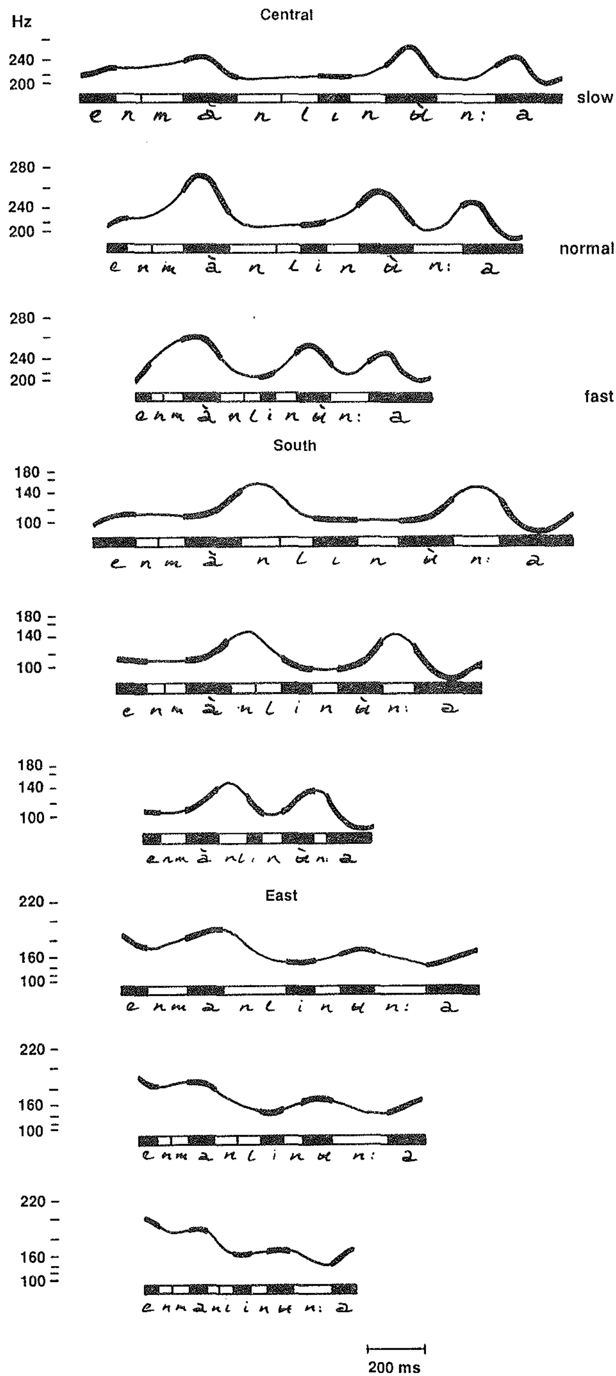


Fig.8. Average durations and F0 curves of 3 repetitions of the sentence en mànli nùnna spoken at three rates by representatives of three Swedish dialects, Central, South, East.

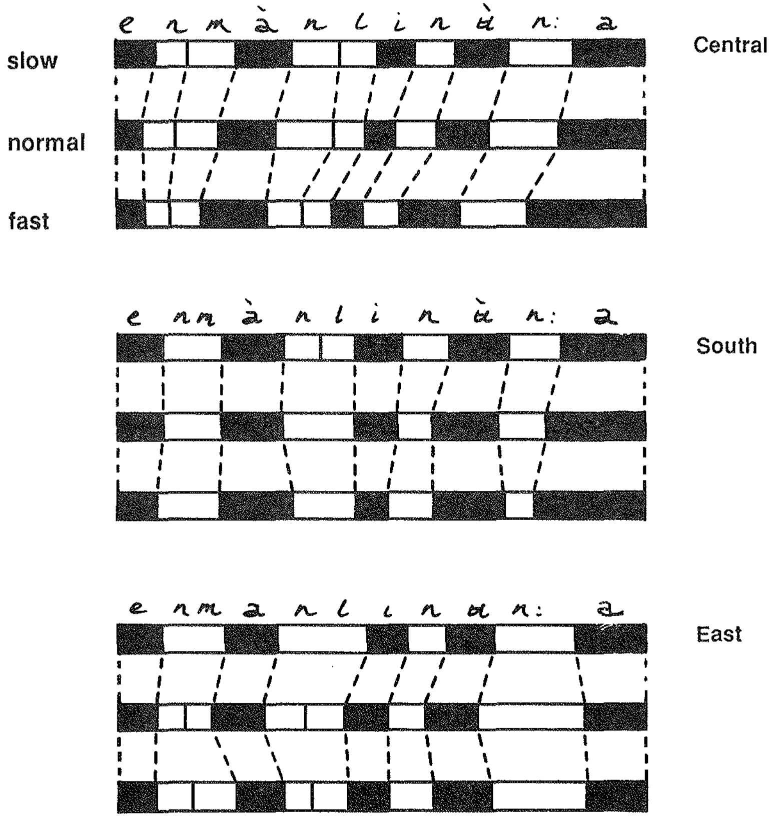


Fig.9. Relative durations of segments in three tempi. Test sentence en mànli nùnna in three Swedish dialects.

dialect the carrier of a sentence accent peak, occupies a much larger part of the fast utterance than it does for the southern speaker who has a very small relative lengthening of the final syllable. The Skåne rhythm (by outsiders judged to sound plodding) retains its characteristics all through the three tempi.

The temporal pattern of the East Swedish speaker is somewhat irregular. However, the first two syllables always occupy more room in the fast utterance than in the slow one, perhaps an indication of an early placement of sentence accent in this dialect. It is a moot question whether this is a trace of an old stage in Swedish prosody or interference from Finnish, which like all Finno-Ugric languages is characterized by an early sentence accent.

Tonal pattern

Figure 10 shows the different-rate tonal patterns imposed on a common temporal base. The already mentioned interdialectal differences are conspicuous. Central Swedish has three strong accent peaks. Southern Swedish has two strong peaks and Eastern Swedish has two weak ones. The first two peaks are connected with the main accents and the third peak of Central Swedish is due to the sentence accent which comes later than the main accent in this dialect (Bruce 1977). Changes with fast tempo can be described as follows (Fig. 11). The up- and down-going movements connected with the accents, the manifestations of the laryngeal gestures, look like triangles which retain their shapes regardless of tempo. When the segments are shortened, some turning points, which will be regarded as the first targets, are very fixed relative to the segments, some move into the neighbouring segments and syllables, while others disappear.

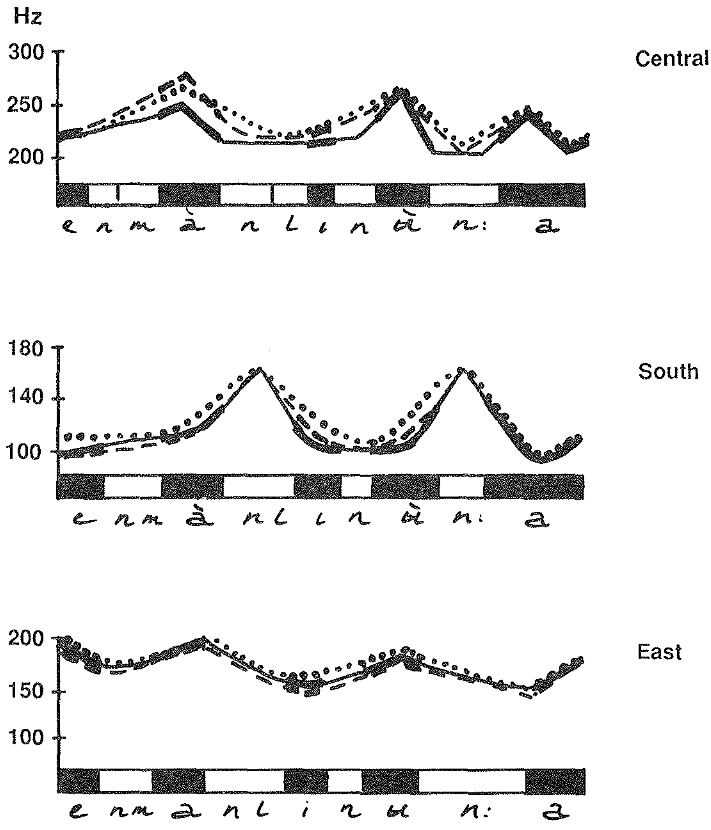


Fig.10. Different-rate tonal patterns imposed on a common temporal base. Test sentence <u>en manli(g) nunna</u> in three Swedish dialects.

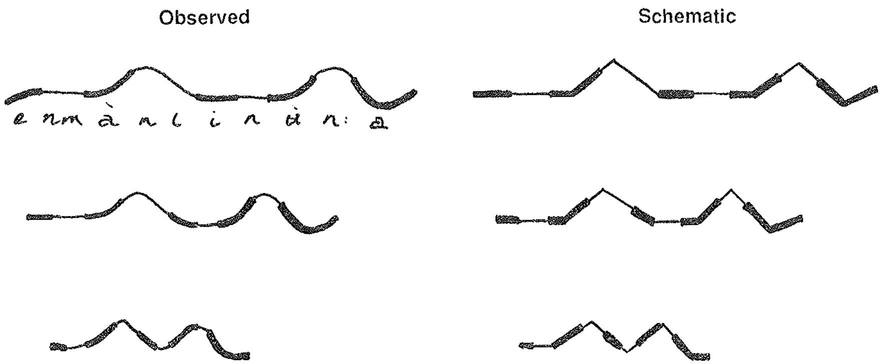


Fig.11. Southern Swedish observed and schematic F0 curves related to segments in three rates of speech. Test sentence <u>en manli(g) nunna</u>. Compare Fig.8.

The figure shows how the flat interval over the consonants and the unaccented syllables of slow speech disappears and is merged in the falling-rising movement of the accents in the fast renderings. It seems, then, that when time is short, it is enough to signal the general direction of the tonal movement over the accented syllable.

For Finland Swedish (see lower part of fig.10), the accent peaks are very weak and the change of the tonal pattern is uniform. The peaks are located at the end of the vocalic segment and the rise starts at the beginning of the preceding consonants regardless of tempo. The remarkable uniformity of East Swedish under tempo changes contrasts strongly with the non-uniformity of Central and South Swedish. The difference may be due to the absence of distinctive accent patterns and is perhaps also due to the influence of Finnish with its syllable-timed prosody.

In the comparison with Chinese, we will concentrate on

Central Swedish as being most similar to Chinese in the behavior of he durational patterns under changes of tempo, a similarity which is probably due to the fact that both these dialects have final sentence prominence.

Laryngeal pattern

In an EMG investigation of speech material focussing on Swedish accents (Gårding et al. 1975) it was observed that at fast rate all the EMG peaks of the vocalis and cricothyroid muscles were slightly higher corresponding to the higher peak values of the pitch curves. Also, the EMG peaks came closer together in fast speech and the cricothyroid peak which at normal rate fell to the baseline of activity turned into a rise well over the baseline. This was correlated to sharp changes from fall to rise in the fast-speech pitch curves as compared to the smooth troughs found in the normal-rate patterns. We notice a similar tendency in the Chinese pitch curves (see e.g. Figures 6 and 7). For Chinese there is also the necessity of adjusting the whole tonal movement to segments in a more precise way than in Swedish which may result in increased steepness and sharper dips in the fast-speech curves.

4. Comparison and conclusions

It follows from the preceding section on Swedish that statements 1-3 about Chinese intonation and the qualifying remarks remain valid for Swedish as well if the word 'tone' is replaced by 'accent'. The reason for the similarity is of course that with increased rate of speech the prosodic commands come closer together with overlapping and weakening of the corresponding gestures as a result. Some examples of this hypothetical process are given in Figure 12.

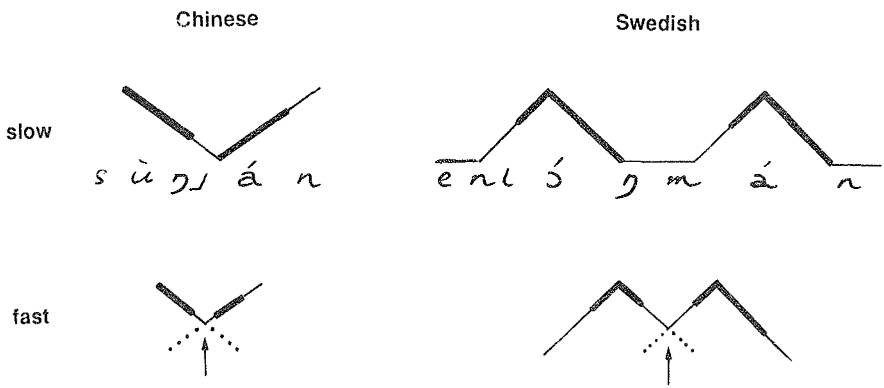


Fig.12. Encounter of laryngeal gestures. Chinese is exemplified by the encounter of a falling and a rising tone, Swedish by two acute (in this dialect falling) accents. Dotted line indicates covert parts of the gestures. Arrows point to the splices. The drawing is consistent with the observed raised F0 values of the low turning points and the narrower grids of fast speech.

What about the observed differences? In our analysis we found that the number of turning points are more numerous and their locations more stable in Chinese than in Swedish. This obviously has to do with the phonological system which results in a larger number of tones than accents in a comparable sentence. In Chinese most of the syllables are tonic and each tone has target turning points tied to the vocalic segment according to certain rules. In spite of sandhi rules and with some allowance for coarticulatory effects the communicative value of a tone will prevent too much overlap. In Swedish an accent is most often tied to a sequence of syllables (the stem) and the number of syllables for which pitch is derived by interpolation is greater than in Chinese. Although the first target point is the most fixed one in both languages there is more room in Swedish for the laryngeal gestures to extend and encompass also neighbouring syllables. As a result, the second target point has more leeway in Swedish than in Chinese.

A phonetic difference which may be worth mentioning and important for the distincion between tone and accent is that the first target point of a Swedish falling accent tends to be located near the middle of the vocalic segment whereas in Chinese a falling tone is more closely connected with the C/V boundary. Another observation is that with increased tempo a Chinese prominent falling tone becomes more similar to its Swedish counterpart in that the peak is shifted towards the middle of the vocalic segment. The larger reduction of less prominent tones in fast speech makes the Chinese fast-speech curves more like those of an accent language.

Finally, let us return to the statement of Kozsenikov and Chistovich that the relative durations of words and syllables are independent of speech rate. In the majority of our dialects, the part under sentence prominence has a longer relative duration in fast speech. Hence their durational patterns are not invariant.

It is also clear from our analysis that, in languages with distinctive tones or accents, the coordination between temporal and tonal patterns has a certain dependence on tempo.

Acknowledgement

We wish to thank the Swedish Academy of Engineering Science for continued support of our collaboration.

We are also indebted to Lars Gårding for the choice of parameters of the grid construction in the appendix.

Appendix

In this appendix we shall outline a way of generating schematic versions of our observed curves using a small number of parameters. The scheme follows in principle Gårding et al. 1983. We assume that the durations of the segments, i.e. the temporal pattern is given according to Figures 6 and 7. The construction does not take microprosody into account.

The outline has three parts, phonological rules, i.e. context-dependent changes of symbols, the tonal grid and the insertion of turning points followed by interpolation.

1. Phonological rules.

Our two sentences have the following sequences of tones in a syntactic structure indicated by parentheses (0 means atonic),

(1) (HL+LH)+((HL)+(LH+HL))#

(2) (HL+HL)+((HL)+(HL+0))#

The sign # marks the end of the sentence. An assimilation rule deletes the second element of every tone when followed by a target of the same kind. We make an exception for the sequence L+L after mai where assimilation does not seem to occur. A sandhi rule weakens a HL between two HL's. The result of these rules is the two sequences where the syntactic structure is left out in the notation

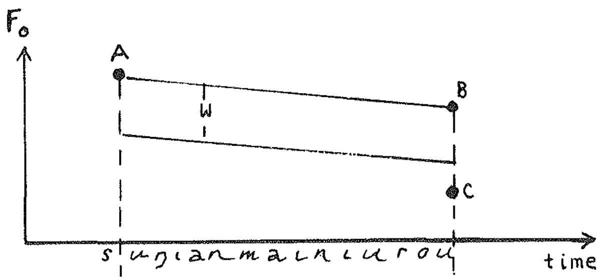
(1) H(L)+L(H)+HL+L(H)+HL#

(2) HL+HL+HL+HL+0#

2. The tonal grid

A falling grid for declarative sentence intonation is constructed using four parameters A,B,C,W. We first describe it as a geometrical figure which shows the role of the parameters. A and

B are the endpoints of a topline stretching over the whole sentence, A being higher than B. The baseline is parallel to the topline at the distance W below it. The point C marks the end of the sentence (#) and the end of the tonal curve. See the figure below.



Average values in Hz of the parameters for three tempi and the two male speakers may be given as follows

A: $195^{+}_{-}10$, B: $165^{+}_{-}10$, C: $90^{+}_{-}10$, W: $45^{+}_{-}5$.

The first figure of each pair refers to normal, + to fast and - to slow tempo. The actual values may differ by 10 or 15 Hz, but for all speakers increased tempo is accompanied by a higher grid and a smaller width. In an extended context including other tonal patterns, it is possible that relative values can be given to the parameters A,B,C and W. To account for different voice registers, the scale should be logarithmic.

3. Construction of the tonal curve

When the grid is drawn for any tempo, the highs and lows which remain after the application of the phonological rules are inserted on the tonal grid above their respective time positions. The weakened highs are inserted halfway between the grid lines. The final step is a linear interpolation between the generated points.

References

Browman C.P. & L. Goldstein. 1986. Towards an articulatory phonology. Phonology Yearbook 3:

Bruce G. 1977. Swedish word accents in sentence perspective. Travaux de l'Institut de Linguistique de Lund XII. Gleerup, Lund

Cheng C.C. 1973. A synchronic phonology of Mandarin Chinese. Mouton, The Hague

Gårding E. 1967. Internal Juncture in Swedish. Travaux de l'Institut de Phonétique de Lund VI. Gleerup, Lund

Gårding E. 1975. The influence of tempo on rhythmic and tonal patterns in three Swedish dialects. Working Papers 12, Department of Linguistics, Lund University: 71-84

Gårding E. 1981. Contrastive Prosody: A model and its application. Studia Linguistica 35: 146-166

Gårding E. 1984. Comparing intonation. Working Papers 27. Department of Linguistics, Lund University: 75-100

Gårding E. 1987. Speech act and tonal pattern in Standard Chinese: Constancy and variation. Phonetica 44: 13-29

Gårding E., O. Fujimura, H. Hirose & Z. Simada. 1974. Laryngeal control of Swedish word accents. Working Papers 10. Department of Linguistics: 53-82

Gårding E., J. Zhang & J.O. Svantesson. 1983. A general model for tone and intonation in Standard Chinese. Working Papers 25. Department of Linguistics, Lund University: 53-65

Gårding E., P. Kratochvil, J.O. Svantesson & J. Zhang. 1986. Tone 4 and Tone 3 discrimination in modern standard Chinese. Language and Speech, Vol.29, Part 3: 281-293

Howie J.M. 1974. On the domain of tone in Mandarin. Phonetica 30, 129-148

Kozsevnikov V.A. & Chistovich L.A. 1965. Rech, artikulyatsiya i vospriyatiye. English translation: Speech: Articulation and perception. Published by U.S. Department of commerce, Washington D.C. 1966

Kratochvil P. 1968. The Chinese Language today. Hutchinson. London

Wu J.-J. 1982. Rules of intonation in Standard Chinese. Preprints from the 13th International Congress of Linguistics. Eds. H. Fujisaki & E. Gårding