Working Papers 25 1983
Linguistics-Phonetics
Lund University, 53-65

A generative model for TONE AND INTONATION IN STANDARD CHINESE based on data from one speaker

Eva Gårding, Zhāng Jiālǔ* and Jan-Olof Svantesson

In most descriptions of Chinese prosody, the basic units are the morphemes with their tone patterns, and these patterns are then modified by sandhi, stress, intonation, and tempo (see e.g. Wū 1982a, 1982b). Figure 1, borrowed from Kratochvíl's well-known manual The Chinese Language today (1968) will serve as an example. In the figure the author shows how a sequence of morphemes, the basic units, develops into a sentence which is supposed to be part of continuous speech. At the first stage we have the individual morphemes and their tone patterns, i.e. from left to right, high, atonic, dipping, dipping, falling, dipping and high tones. The second stage shows the results of the sandhi rules. For instance, in a sequence of two dipping tones, the first is turned into a rising tone. At the next three stages, the effect of stress is exemplified, e.g. a reduction of duration and tone in the pronoun wǒ 'I', and the quantifier běn 'a volume of'. The last stage shows the workings of sentence intonation. The most conspicuous effect in this case is a fall over the last syllable of the sentence.

What we will do here is to show how we have tried to handle all these phenomena within the framework of a generative intonation model developed earlier for Swedish (Bruce and Gårding 1978) and later expanded to account for the intonation of other languages as well (Gårding 1981, 1983). The main feature of this model is that it separates lexical prosody from phrase and sentence prosody. Table 1 shows examples of the sentences we have worked with. In Figure 2, two examples of the intonation

* Institute of Acoustics, Academia Sinica, Beijing

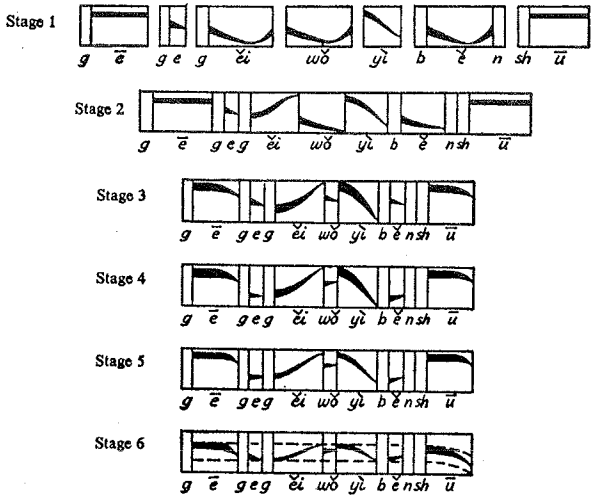


Fig. 17 Modifications of suprasegmental features in the sentence *gēge gěi wǒ yì běn shū*

Figure 1. From Kratochvíl 1968.

Table 1. Examples of test material.

STATEMENTS

I. Focus free	1,2 :	Wāng Yī chōu xiāngyān. Wāng Yī smokes cigarettes.
	3,4 :	Sòng Yán mài niúròu. Sòng Yán sells calf meat.
	5,6 :	Wāng Lǐ chuān yǔyī. Wāng Lǐ wears a raincoat.
II. Focus left	1,2 :	Shì, Wāng Yī chōu xiāngyān.
	3,4 :	Shì, Sòng Yán mài niúròu.
	5,6 :	Shì, Wāng Lǐ chuān yǔyī.
III. Focus right	1,2 :	Wāng Yī chōu xiāngyān.
	3,4 :	Sòng Yán mài niúròu.
	5,6 :	Wāng Lǐ chuān yǔyī.

QUESTIONS

Focus free	1,2 :	Wāng Yī chōu xiāngyān?
	3,4 :	Sòng Yán mài niúròu?
	5,6 :	Wāng Lǐ chuān yǔyī?

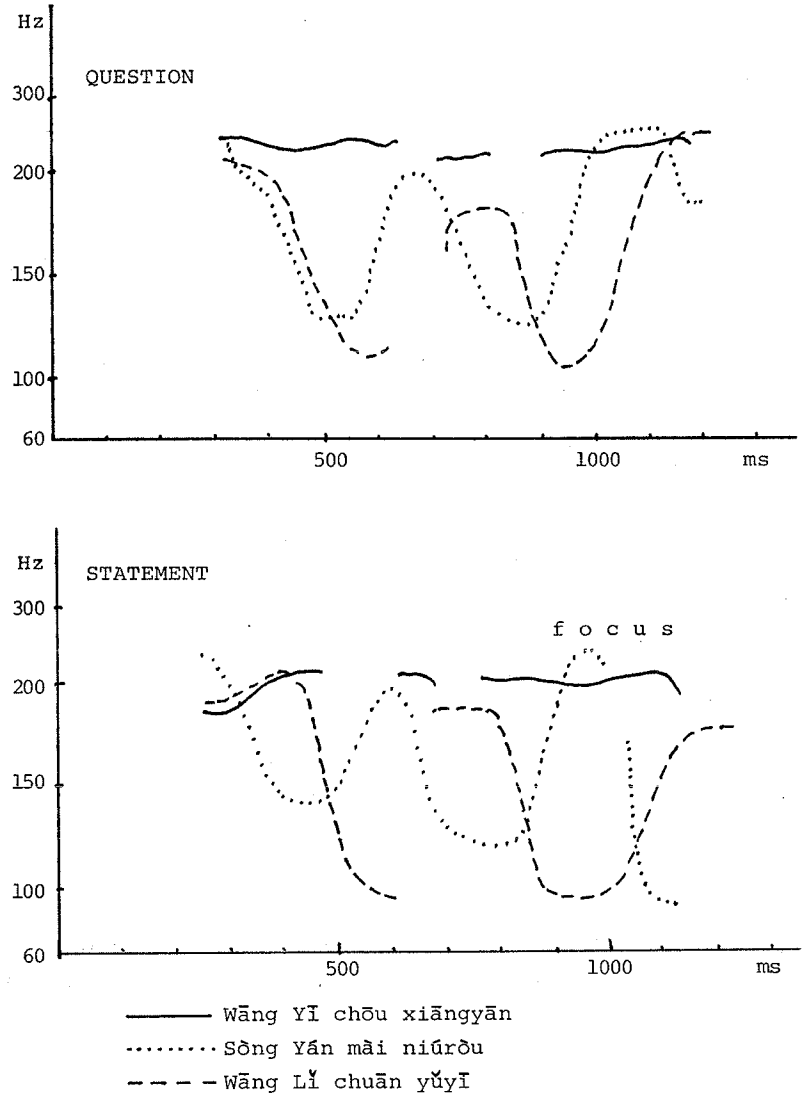


Figure 2. Three sentences with different tone patterns under two different intonations. Focus-free question above, statement with focus on the last word below.

contours are given, each with three different tone patterns.

Figure 3 shows the main concepts and principles of the model. It is here demonstrated how a typical intonation contour has maxima and minima, turning points, which are part of a global pattern, the grid, which can be rising, falling or level. The grid can be expanded or compressed. The part of the grid where its range or direction changes is called a pivot.

The local turning points, as well as the direction and range of the grid and the pivots are associated with communicative events in the following way:

Local turning points	lexical tones, accents (words, morphemes)
Pivots	syntactic boundaries
Direction of grid	speech act (statement, question, etc.)
Range of grid	prominence (focus, out of focus)

When a pitch curve is generated by our model, the grid, that is the global expression of sentence intonation, is generated first. Then the local maxima and minima pertaining to lexical items are inserted as points into the grid according to specific rules which state how the points are aligned relative to the segments. The fact that the location of these points is practically independent of sentence intonation and tempo makes this arrangement natural. In the last step, the pitch curve is obtained by smooth interpolation between the points over the voiced segments.

Figure 4 shows the different stages of the model of which the pitch algorithm is the final part. The stages are the same as for the other languages we have investigated (cf. Gårding 1981). The input is a sentence in which the morphemes are equipped with markings for tones, phrase accent and sentence accent, syntactic boundaries and speech act. It may also be marked for focus. The first three stages of the model take care of the duration and the last three of pitch. So far, this is a practical arrangement which was motivated by the fact that, in an accent language which was behind the original model, syllables could carry accentuation by means of duration only, without accentual pitch movements. The order of the stages for a tone

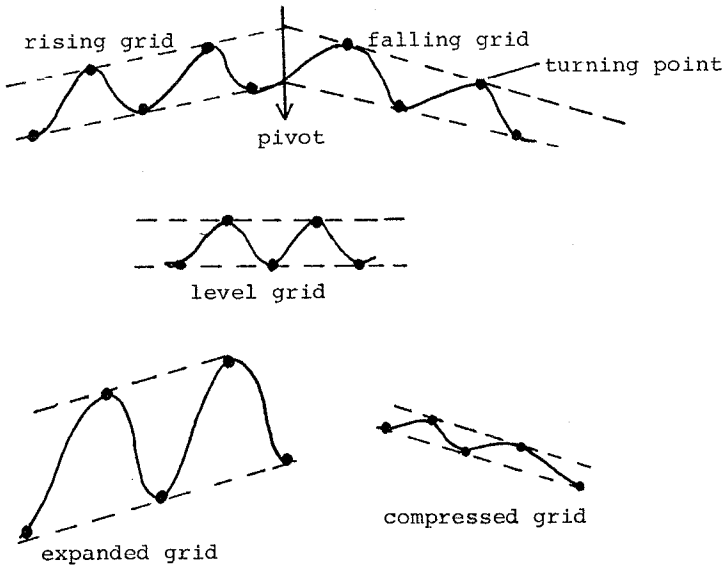


Figure 3. Concepts of the model

---

language must be regarded as very tentative.

We shall here concentrate on the pitch generating part, of which Figure 5 is a close-up. Among the intermediate phonological rules for Chinese we have the sandhi rules, e.g.

T3 ⟶ T2 / ___ T3

which means that a dipping tone becomes a rising tone before another dipping tone.

We also have tone reduction rules. For instance, in a phrase which consists of three morphemes (such as jiàoyùbù 'ministry of education' in the example sentence on Figure 4), the middle one may be reduced in duration and become atonic.

At the next level, the intermediate pitch representations, the abstract tone symbols are converted into more concrete ones, Lows and Highs, and combinations of Lows and Highs. These symbols are rough approximations of the tone contours of the dialect we are analyzing. It should be noted that they are not

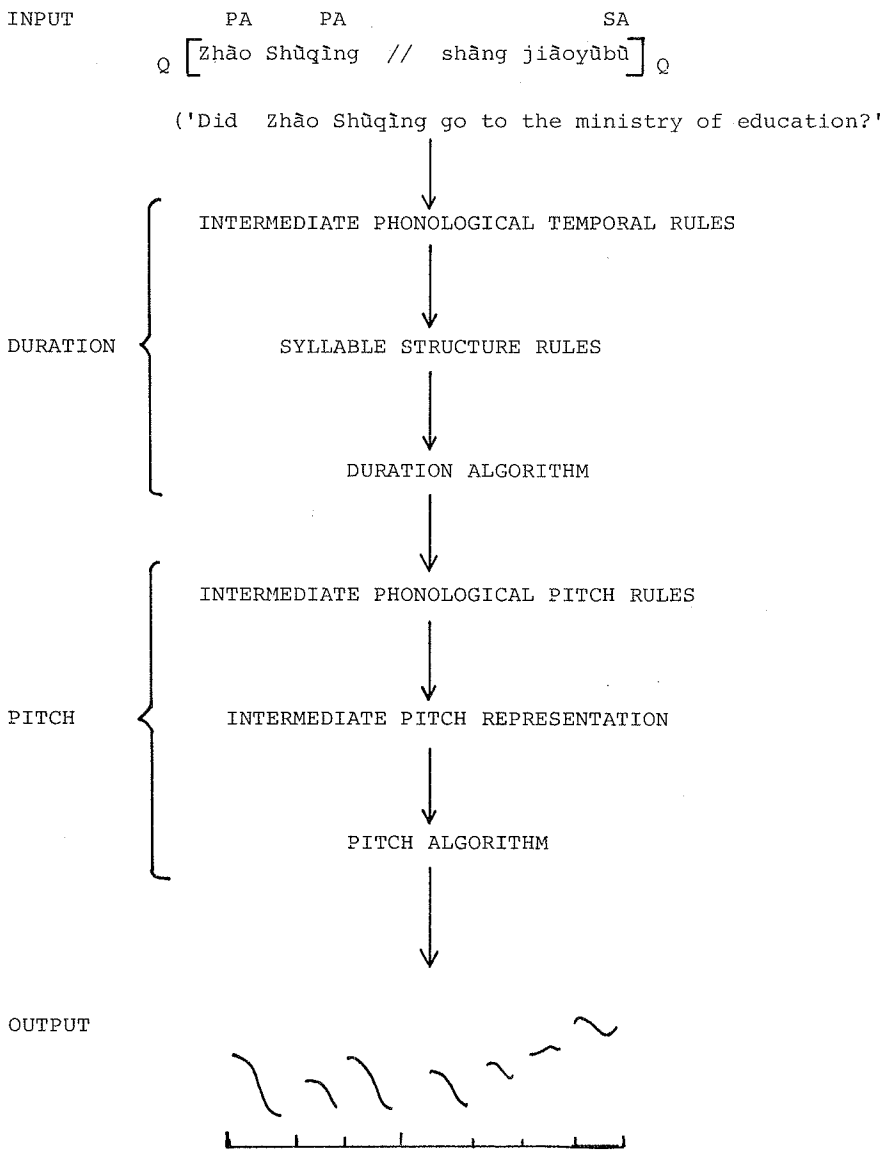


Figure 4. Model for prosody.

Intermediate phonological rules (symbol level)

Sandhi rules, e.g. T3→T2 / ___ T3
Tone reduction rules

Intermediate pitch representations (symbol level→more concrete level)

Tones: T1→HH
T2→LH
T3→LL
T4→HL

Speech acts:
Statement: Global fall and/or local fall
Question: Global rise and/or local rise
Focus: Part in focus expanded and/or
Part out of focus compressed

Boundaries: Pivots of various kinds
Time position predictable from syntactic structure mainly

Pitch algorithm (cf. Figure 6)

Figure 5. Close-up of the pitch-generating part of the model.

---

equivalent to citation forms. To obtain a citation form, one of of these tone symbols has to pass through the pitch algorithm with a grid expressing statement.

Markings for speech act and boundaries are turned into concrete representations bearing on the grid:

- A statement has a global fall over the final phrase
- A question has a global rise over the final phrase
- Focus is represented by an expanded grid
- Out of focus is represented by a compressed grid
- The boundaries are converted into pivots

All these global features seem to be next to universal.

The final stage is the pitch algorithm (Figure 6). The first rule concerns the grid, i.e. the global frame for sentence intonation. This grid is drawn using information about speech act and syntax.

The second rule inserts the tone marks.

The third rule adjusts the tone marks according to the context.

The last rule takes care of interpolation.

Attached to the pitch algorithm there are prescriptions for determining the phonetic values needed to generate the output signal. These are exemplified in Figure 7.

First we need to know the relative levels that the pitch moves between, the floor (L) and the ceiling (H) of the normal voice register, and the levels of the highs (h) and the lows (l). We call this scheme of levels the stave. For our speaker, the range between the floor and the ceiling is about one and a half octaves, and the range between high and low tones one octave. Visual inspection of the intonation contours from our other informants shows that this is not an individual characteristic. Our Swedish speakers have a range of about an octave. For our Chinese speaker, the range narrows in fast, informal speech.

We also need prescriptions for the grid. In the exemplified case, which is a question, we have a pivot in the middle, a grid rising to the ceiling at the end of the final phrase, and a falling grid for the preceding phrase.

We need to know how the grid is expanded and compressed for the part which is in, or out of focus.

Only one example of an alignment rule will be given here, the one for the dipping tone, T3. It says: reach level L in the middle of the vocalic segment and stay there for a short interval. This lingering is behind our representation, LL. It is the only tone that always comes down to this level and it seems to be a necessary condition, as shown informally by experiments using the LPC technique.

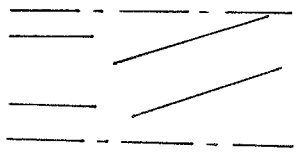
There are at least three kinds of context rules. They express undershoot (assimilation), sharpening of a contrast (dissimilation), and priority given when there is a clash between commands at phrase level and lexical level.

PA PA SA

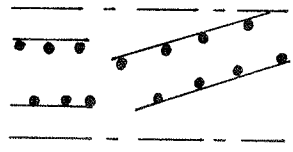
INPUT [Zhāo Shūqìng / shāng jiàoyùbù?]

RULES

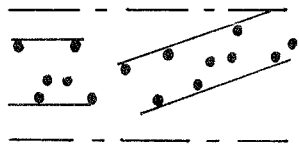
1. Draw the tonal grid:



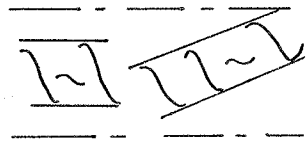
2. Insert the tone marks:



3. Context rules:



4. Interpolation and smoothing:



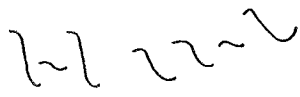
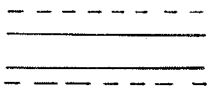
OUTPUT 

Figure 6. Application of the pitch algorithm

Levels in the register: the stave



How to draw a grid

Alignment rules for tones in relation to segments

Context rules: Assimilation
Sharpening

Smooth interpolation

Figure 7. Prescriptions and conventions of pitch algorithm

We shall give one example of each:

Undershoot: in a sequence consisting of high tones, these tones tend to undershoot Level h.

Sharpening: in a sequence of high tones, there are, particularly in slow speech, small dips between the highs as if to emphasize the highs. There is a corresponding sharpening of the lows in a sequence of low (dipping) tones, by introducing rise-falls. Only in this case, the effect is constant and has a larger range. It is a phonological sandhi rule (T3 $\longrightarrow$ T2 / ___ T3), and has its place among the intermediate phonological rules.

Priority: the constant feature of focus is an expanded pitch range. The frequency position of this range is determined by the tone. For a rising tone, the ceiling will be given priority to achieve the expansion, and for a falling tone the floor.

Interpolation: cosine functions may be used for the interpolation between the generated points.

The earlier shown figure, Figure 6, shows the workings of the four rules for a particular input sentence.

Figure 8 gives an idea of the algorithm generating our demonstration sentences of Figure 2. The existence of the pivot is strengthened by the contour of the long sentence below (Figure 9).

Some well-known physiological and psychological principles are at work in the rules of the model:

The inertia principle is reflected in the context rules which take into account the difficulty of large movements over a short interval of time.

The least effort principle is also apparent in the context rules in various short-cuts of F_0-movements which do not lead to an impairment of communication.

The contrast principle may be at work in the rules of sharpening and in the rules of priority.

Finally, the look-ahead principle finds its expression in the grid.

This broad-brush analysis of one speaker's Standard Chinese intonation will serve as a model for the treatment of our remaining data, collected from five more speakers of Standard Chinese.

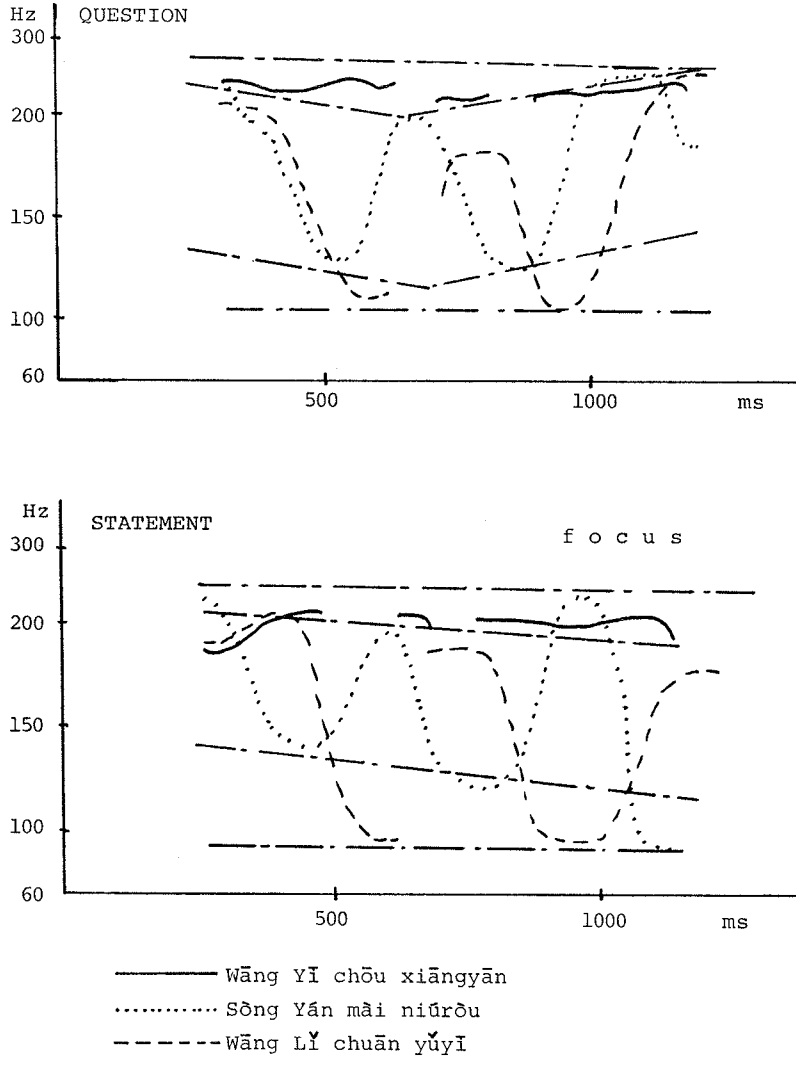


Figure 8. Grid superimposed on intonation contours of Fig. 2

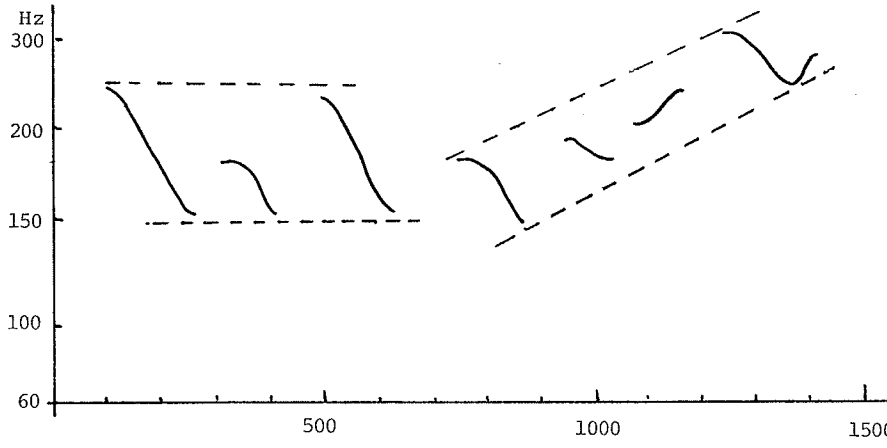


Figure 9. Pivot, illustrated by the sentence Zhào Shùqìng shàng jiàoyùbù? 'Did Zhào Shùqìng go to the ministry of education?'

REFERENCES

Bruce, G and E Gårding. 1978. A prosodic typology for Swedish dialects. In Gårding, E., G Bruce and R Bannert (eds.) Nordic prosody. Travaux de l'Institut de Linguistique de Lund 13, pp. 219-28. Lund:Gleerup

Gårding, E. 1981. Contrastive prosody: a model and its application. Studia Linguistica 35,1-2:146-65.

Gårding, E. 1983. A generative model of intonation. In Cutler & Ladd (eds),Models and measurements. **Spr**inger. In press.

Kratochvíl, P. 1968. The Chinese language today. London.

Wú, Z. 1982a. Pǔtōnghuà yǔjù zhōng de shēngdiào biànhuà (Tone changes in Standard Chinese sentences). Zhōngguó yǔwén 1982:6,439-50.

Wú, Z. 1982b. Rules of intonation in Standard Chinese. Preprints from the 13th International Congress of Linguistics, Tokyo.

This paper was presented by Gårding at a conference, Prosody normal and abnormal, arranged by the Association Européenne de Psycholinguistique in Zürich.

Our work is part of a project, Fonetisk beskrivning av några viktiga språk i tredje världen, sponsored by the Research Council for Humanistic and Social Sciences.