Pitch Target of Neutral Tone in Standard Chinese

Yiya Chen[1] & Yi Xu[2]
[1]Edinburgh University, yiya@ling.ed.ac.uk; [2]Haskins Laboratory, xu@haskins.yale.edu

Abstract: It is commonly assumed that distinctive tonal features are realized as F_0 turning points. Specifically, an F_0 peak is the phonetic realization of a phonologically specified high tone while an F_0 valley is that of a low tone. Tonal Targets are interpolated which gives rise to a continuous F0 contour as well as surface F_0 values of linguistic units that are not specified with underlying tonal targets. We re-evaluate such a view by investigating the F_0 realization of neutral tone syllables in Standard Chinese, which are widely accepted as toneless underlyingly. The results of our experiment indicate that neutral tone syllables do have a target that is independent of the surrounding tones, although the surface realization of this target is dependent on various factors. Furthermore, the neutral tone seems to be different from full lexical tones in the manner with which its tonal target is implemented: it is slow both in overcoming the influence of the preceding full tone and in approaching its own target. Applying the recently proposed pitch target approximation model, we conclude that neutral tone differs from other lexical tones only in the articulatory strength with which tonal targets are implemented, and that articulatory strength constitutes an important coding scheme in speech.

Background

- **Phonetic Implementation of Phonological Tonal Targets**
  - ➡ It is commonly assumed that distinctive tonal features are realized as F_0 turning points. Specifically, an F_0 peak is the phonetic realization of a phonologically specified high tone while an F_0 valley is that of a low tone.
  - ➡ Targets are interpolated which gives a continuous F_0 contour. (Pierrehumbert 1980)
- **F_0 Realizations in Standard Chinese**
  - ➡ Most syllables are specified with underlying lexical tones and surface with distinctive F_0 contours.
  - ➡ There are syllables, called neutral tone syllables, that exhibit rich contextual F_0 variations.

Basic Tonal Inventory: Four distinctive lexical tones

1st tone:	High [H]	ma 'mother'	2nd tone:	Rising [R]	ma 'hemp'
3rd tone:	Low [L/LH']	ma 'horse'	4th tone:	Falling [F]	ma 'scold'

Neutral Tone Distribution

a. Grammatical morphemes (which never surface with any of the four lexical tones)	chi[H] *le*	'have eaten'
b. Reduplication	mei [F] *mei*	'sister (diminutive)'
c. Bi-syllabic lexical items	li [R] *ba*	'fence'

(Note: Neutral tone syllables are in italics.)

Received wisdom: Neutral tone syllables are not specified with underlying tonal target.

Debate: Surface F_0 realization

1) ***Spreading*** from the preceding lexical tone (Yip 1980)

2) ***Interpolation***

2a) between preceding and following lexical tones (Shih 1987)

2b) between lexical and intonational boundary tones (Li 2003)

• **Goals of the Study**

Þ To seek experimental evidence for the phonological and phonetic nature of the F_0 contours over neutral tone syllables

Þ To test the received view of target and target interpolation

Method

Test Material

Sentence medial neutral tone sequences. 32 sentences varying in

1) Tone of preceding syllable (4 lexical tones)

2) Tone of following syllable (2: falling, low (rising)).

3) Number of neutral-tone syllables (0 to 3)

e.g. Ta shuo **(ma [H] *ma men de)*man[F]** duo le.
He said mother pl. poss. slow more asp.
'He said that the mothers' ones are much slower.'

Procedure

- *Subjects*: 2 male and 2 female speakers of Standard Chinese
- *Renditions:* On-focus and pre-focus; fast and normal speaking rates
- *Repetition*: 3

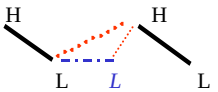
Measurement

- Segment labeling and F_0 extraction in Praat
- Vocal pulsing and segment files are processed with a set of C programs to obtain smooth and time-normalized F_0 contours (Xu 1999)
- Peak velocity of F_0 of the last quarter of the syllables that carry the first neutral tone or the following lexical tone
- F_0 excursion size of the first neutral tone syllable or the following lexical tone when there is no neutral tone, in semitone (= 12 * log2(F0 / $F0_{reference}$ the F0 value of the end point of the syllable *shuo* 'speak').
- Stiffness of the F_0 movement (stiffness = peak velocity (st/s) / excursion size (st))

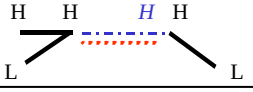
Predictions (1): Spreading vs. Interpolation of Lexical Tones

* Different schematic pitch contours predicted by **spreading** and interpolation

ma (F) + Neutral + man (F)



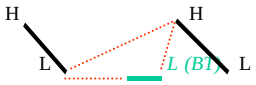
* **Both predict: ma (H/R) + Neutral + man (F)**



Predictions (2): Interpolation of Lexical vs. Intonational Tones

* Different schematic pitch contours predicted by interpolation between lexical tones and interpolation between lexical and boundary tones

ma (F) + Neutral + man (F)

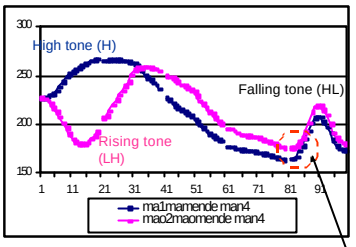


* **Both predict**:

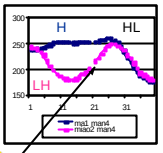
ma (F)+ Neutral + BT + mei (L)

7

Results (1): Neither Spreading nor Interpolation of Lexical tones



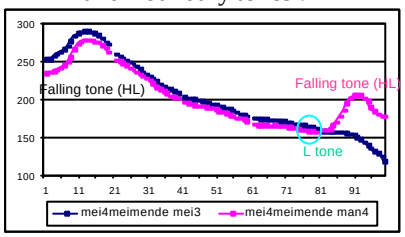
No Neutral Tone Syllable: lexical tones (H/LH) followed by lexical tone (HL)



Three neutral tone syllables preceded by lexical tones (H/LH) and followed by lexical tone (HL)

8

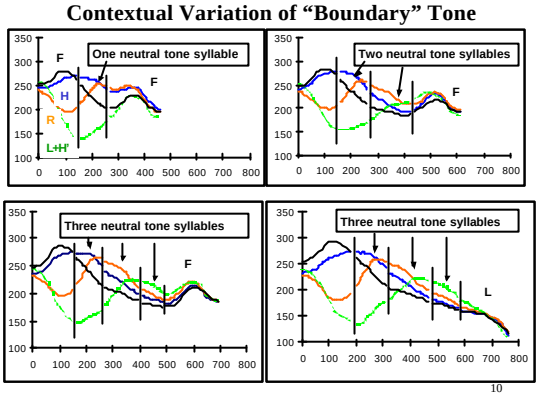
Results (2): Interpolation between Lexical and Boundary tones ?



Question: How to explain the wide range of contextual variations of the "boundary" tone, which is not seen on a lexical tone?

9

Contextual Variation of "Boundary" Tone



10

Two existing models of contextual variations

A. Target Connection: Targets are invariantly specified for an F_0 value, various ways of connection (Browman & Goldstein 1992)

B. Target Specification: Targets may be specified for a wide range of F_0 values. (Keating 1990)

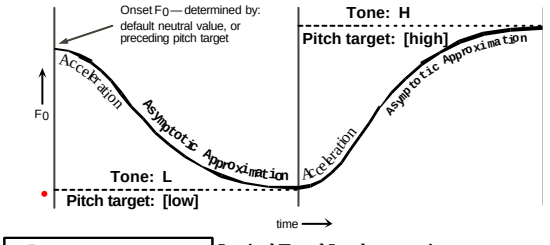
But: Neither explains

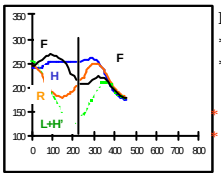
1) Why the surface F_0 contours over neutral tone syllables becomes less varied as the number of neutral tone syllable increases.
2) Why there is a lingering effect from the preceding lexical tone even on the 3rd neutral tone syllable.

An alternative account: Target approximation

The production of a tone is a process of continually approximating its underlying target within allocated time. The approximation terminates when the allocated time is over (Xu & Wang 2001).

11





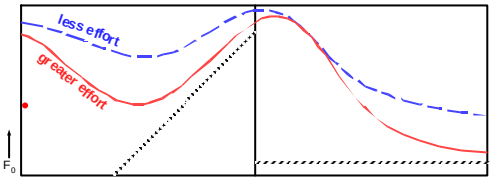
Lexical Tonal Implementation

* Implemented synchronously with the syllable.
* Approximated near the right edge of the tone carrying syllable.

* This would predict variability with preceding tone
* But the tonal target over a neutral tone syllable does not seem to be reached within one syllable

12

Effect of reduced articulatory effort



Supporting evidence from segment production

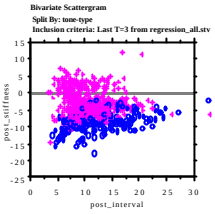
Perkell et al. (2002a, 2002b) show that clear speech is produced with greater articulatory effort than normal speech

Hypothesis: Lexical tones are implemented with greater articulatory effort than neutral tone.

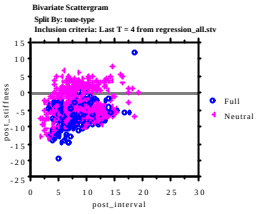
Prediction: Peak velocity of the F_0 movement over a lexical tone is greater than that of a neutral tone as peak velocity is a good indicator of the effort that speakers exert to produce an articulatory movement (Nelson 1983, Perkell et al., 2002a & b)

1) L, F and neutral tone all happen to have declining f0 contours.
2) At various pitch excursion sizes, peak velocity and stiffness of L and F have greater peak velocities than the neutral tone.

Full Tone is 3rd Tone (L)



Full Tone is 4th Tone (F)



Stiffness = peak velocity (st/s) / excursion size (st)

(semitone = 12 * log2(F0 / $F0_{reference}$))

14

To conclude

- The most coherent account for the varied F0 contours of the neutral tone tone in Mandarin is to assume that it has a pitch target of its own.
- Like that of a full tone, the pitch target of the neutral tone is continually approached within its host syllable.
- Unlike that of a full tone, the pitch target of a neutral tone is implemented with much less effort.

This explains

. * Lingering influence of the preceding lexical tone over a span of neutral tone syllables

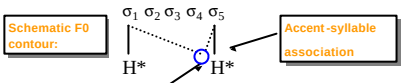
* Converging reports in the literature that the F0 contour of neutral tone is determined by the preceding lexical tone

15

Implications for Intonational Phonology

English Intonation (Pierrehumbert 1980)
Tonal targets e.g. L+H*, H*
Intonation contour: Linear interpolation of tonal targets.

A Problematic Case: H* H* sequence shows declining pattern.



Solution: "Sagging transition"

New finding: (*Ladd and Schepman, 2003*)

Both production and perception studies indicate that there is a L target in the original sagging transition.

16

Remaining problem for the existing theory:

How to account for the differences between the L in H*-(L)-H* and the L in H*-L+H* tones if both have an L target?

Observations to be noted:

1. In H*-(L)-H*, the valley is much higher when there is no intervening unaccented syllable.
   -> *Comparable to one-neutral-tone sequence.*
2. Declining pitch contour between the sequence of H*-(L)-H* tone as the number of unaccented syllables increases.
   -> *Comparable to longer neutral-tone sequence.*

One possibility:

The L target in the H*-(L)-H* tone has a tonal target similar to the pitch target observed over neutral tone syllables in Standard Chinese.

17

Selected References

Browman, C. & L. Goldstein. (1992). Targetless schwa: an articulatory analysis. Papers in Laboratory Phonology II: Gesture, segment, prosody. G. D. & R. Ladd, Cambridge University Press: 26-36.

Keating, P. (1990). The window model of coarticulation: Articulatory evidence. Papers in Laboratory Phonology I: between the grammar and physics of speech. J. M. B. Kingston. Cambridge, Cambridge University Press.

Ladd, R. & A. Shepman (2003) "Sagging transitions" between high accent peaks in English: experimental evidence. Journal of Phonetics.

Li, Zhiqiang. (2003). The Phonetics and Phonology of Tone Mapping in a Constraint-Based Approach. PhD dissertation, MIT.

Nelson, W. L. (1983). Physical principles for economies of skilled movements. Biological Cybernetics 46: 135-147.

Perkell, J. S., M. Zandipour, M. L. Matthies and H. Lane (2002a). "Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues," J. Acoust. Soc. Am. 112, 1627-1641.

Perkell, J. S. and M. Zandipour (2002b). "Economy of effort in different speaking conditions. II. Kinematic performance spaces for cyclical and speech movements," J. Acoust. Soc. Am. 112, 1642-1651.

Pierrehumbert, J. (1980). The phonology and phonetics of English intonation. Cambridge, MIT.

Shih, C. (1987). The phonetics of the Chinese tonal system. AT&T Bell Labs technical memo.

Xu, Y. (1998). Consistency of Tone-Syllable Alignment across Different Syllable Structures and Speaking Rates. Phonetica 55(4): 179-203.

Xu, Y. & Q. E. Wang (2001a). Pitch targets and their realization: Evidence from Mandarin Chinese. Speech Communication 33: 319-337.

Yip, M. (1980). The tonal phonology of Chinese. Cambridge, MIT Press.

18