# Prosodic Diversity in Mandarin Natural Discourse

Li-chiung Yang
Department of Linguistics, Georgetown University
Washington, D.C. 20057 U.S.A.

## ABSTRACT

*Expressive rhythm and melodic structures permeate natural discourse. This study examines the prosodic system of Mandarin Chinese by acoustically and perceptually analyzing discourse data. Pitch shapes often reflect discourse intentions, underlying cognitive states, and accommodation to intonational conflict and progression. It is found that rhythm, iconicity and inversion act fundamentally to connect sound with meaning.*

## INTRODUCTION

Intonation ultimately depends on the sympathetic unity of emotions which are called up by specific sound characteristics. Intonation mirrors the complexity of cognitive and emotional states throughout a dialogue, and small waves of finely focused micro-emotions simultaneously exist within larger waves of more broadly focused emotions. The complex cognitive and emotional layers are reflected in a rapidly changing process of intonational progression. For example, the arched rise fall shape of dawning realization iconically represents the transformation from surprise and uncertainty to certainty and acceptance. The progression of intonation parallels the progression of emotion.

## SHAPES OF INTONATION

Pitch shapes of wiggles, twists, turns and waviness are characteristic of animated and involved speech, and often reflect discourse intentions, underlying cognitive states, and accommodation to intonational conflict and progression. Wiggles, upturns and downturns also contribute to a feeling of pleasure and agreeableness. Twists in pitch shape can arise from the strain to accommodate divergent emotions. Smooth wiggles and waves may result from a more relaxed state of the vocal cords and twists from tension of the vocal cords.

Micro pitch turns are also used to resolve the conflict between intonation and lexical tone. Counterbalancing micro-turns at the end of syllables can also enhance rhythm. A tiny pitch reversal at the end of a syllable often indicates either an emotional state whose intonational shape is opposite to the predominant pitch shape of the lexical tone, or a tiny indication of lexical tone attached to an opposite intonational shape of the main part of the syllable. The pitch shape of a rising second tone syllable in an environment of rejection or negativity often takes on a falling shape with a slight pitch upturn at the syllable end. The extreme sensitivity of the ear and brain in obtaining critical information from these micro-turns emphasizes the importance of even slight pitch movements in intonation.

## DISCOURSE INTONATIONAL RHYTHM

Rhythm is established by the repetition of specific aspects of sound or sound patterning at regular time intervals. Rhythmic repetition can be inexact, as in a natural progression

of vowel sounds from open to closed, or vice versa. Rhythm can also be evoked by the timewise balance of natural sound inverses, such as high vs low, strong vs weak, sharp vs gradual, and quick vs slow.

Poetic elements are pervasive in natural discourse and are manifested through repetition, sound harmony and melodic pattern variations. Reduplication can be seen in the example "wo3 kan4 zhao3 zhao3 kan4" meaning "I'll look to see". There is a great contrast in the pitch rise of the first *kan* and the pitch fall of the second *kan*. The first *kan* is self-questioning and tentative, and so rises. By the second *kan*, the speaker has already decided on a course of action, is more definite and certain, and so the pitch falls. The second *zhao3* of the reduplicated form *zhao3 zhao3* psychologically completes the reduplication and so has a short fall. Such intonational couplets of rising and falling form often accompany proposition and completion pragmatic structures. This association of fall and low with completion and concluding may also explain why unit final third tones are often realized as a simple fall and may partially explain why tone sandhi (33-->23) occurs in speech.

Reduplication makes things more casual and lighter. Reduplication, repetition and rhythmic speech often serve to increase familiarity and intimacy. This may be because by using rhythm, the speaker adds a pleasant sound sensation to the ear beyond the purely semantic meaning. Another possible reason for reduplication is that it allows one of the syllables to take on greater intonational variation. *Intonational* repetition occurs frequently in speech. In the example of Figure 1 "London hen3 re4, jiu3 shi2 du4" [London's very hot, ninety degrees], the striking and exaggerated intonational pitch pattern of *London hen re* is repeated in the phrase *jiu shi du*. The reasons for such patterns of repetition are that an expression of emotion can be emphasized or prolonged by repetition of the physiological speech sensation associated with an initial emotion. Time spent in sensory perception may also increase the cognitive impression.

Rhythm connects. Through assignment of equal or hierarchical acoustic forms, syllables, words, or phrases achieve equal or hierarchical status. In this way, rhythm can be used for organizational effect, and a hierarchy of grammar and meaning can be achieved. For example, a frequent phrase or sentence pattern found in my data is to first establish a specific rhythm, and then to conclude by breaking the rhythm, sometimes by using an inverted rhythmic form. In the example [you4 zhun3] [wai4 guo2] [xue2 sheng1] [shu3 jia4] [han2 jia4] [ke3yi3 da3 gong1], "allow foreign students to work during summer and winter vacations", syllable couplets were formed by rhythmic timing through *han jia*, then the rhythm was broken to conclude by the timewise merging of *ke* and *yi* and the lengthening of *gong*. In another example the phrase "yi1 bai3 duo1 ge0 zhong1 guo2 xue2 sheng1" [over one hundred Chinese students] was said with a H-L H-L H-L L-L pitch level sequence, concluding with two lows. Establishing a rhythm and then breaking it performs a cognitive or discourse function, and the break serves to signal the conclusion of the rhythmic unit. The notion of rhythm is fundamentally related to cognitive signal-processing efficiency. Our cognition and senses may be more attuned to signals of a periodic nature.

Rhythm is enhanced by symmetry. In "kou3 shi4 ye3 hai2 mei2 kao3, lun4 wen2 hai2 mei2 xie3" [Oral exam I still haven't taken, dissertation I still haven't written], the phrases *kou shi ye hai mei kao* and *lun wen hai mei xie* are rhythmically paired by the equal amount of time spent on each phrase and their parallel syntactic construction. The 2nd half of the phrase is slightly lower in pitch and amplitude for completion, but there is no downdrift. That preserves the poetic equality.

In discourse, there is often a rhythmic acceleration as one works away from the point of most emphasis. In the example "bu2 guo4 ta1 li3 mian4 you3 san1 shi2 ji3 ge0 college" [but inside they have thirty some colleges], there is strong emphasis on *san shi*. Within *san shi ji ge*, *san* carries the main focus and therefore is the longest, and the subsequent syllables are progressively shorter. The speech tempo is often iconic in that we dwell on things which are more important or pleasant, and spend less time on deemphasised points.

Transitions between rhythm are expressive. In the example of Figure 2 "qi1 nian2 nian4 bu2 nian4 de0 wan2 bu4 zhi1 dao4" [Whether or not I can finish studying in seven years, I don't know], the phrase *qi nian nian bu nian de wan* is said as a smooth descending pitch unit and then changes to a short and choppy *bu zhi dao*, with downward pitch force in each syllable. Negativity is signalled by the repetitive downward pitch falls of *bu zhi dao*. Choppy or clipped speech in general seems to iconically represent a reluctance to speak and was often associated with embarrassment, unhappiness, or complaining.

In the example "bu2 guo4 zhi4 shao3 ni3 gan3 jue2 zai4 mei3 guo2 zhi4 shao3" [But at least you feel that in America, at least], an initial dramatic disagreement is signalled by the very high pitch and strong amplitude on *zhi shao*. By repeating and normalizing the same phrase *zhi shao* at the end, with lower amplitude and pitch, the speaker is making a more convincing case. The phrase *zai mei guo* sounds like intonational repetition of *ni gan jue*, and the particular repeated pitch pattern of the fall rise is very expressive of "trying to convince" in this example.

## ICONICITY AND INVERSION

The notion of iconicity is one which can unify many aspects of intonation. There are different levels of iconicity, and sounds reflect iconicity to different degrees. The data reveal that regular associations between sounds and other non-sound entities exist. Iconicity at one level is connected with the body, providing speech examples such as throat constrictions and high pitch levels. Beyond that, speech sounds are associated to non-sound entities in a metaphoric way or by parallelism, as in sound symbolism. Iconic elements often seem to reflect the cognitive nature of the mind and the perceptual demands of physiology, e.g. in repetition or in lengthening of duration for importance, and iconic elements are related to emotion: we dwell on things we like, we shorten and disregard things we don't like.

Iconicity leads directly to the *principle of inversion*. If speech is iconic, if sounds iconically represent mental and physical states, then when one wants to express an opposite emotion or state, one uses the opposite sound. Sound opposites include high vs low pitch, strong vs weak amplitude, quick vs slow tempo, long vs short duration, smooth vs non-smooth rhythm, rising vs falling pitch, etc. In discourse, the flow of changing conditions and states may often call for a need to overcome and reverse the preceding intonational effect and may result in intonational flip-flop, that is, an inverse intonation may be used just to counteract the just preceding unit.

Why should inversion be so important in speech? In natural speech, discourse structures of climax and resolution or intensification and normalization appear constantly, both on a small scale of a few syllables and on a large scale across phrases or sentences, as in stories or episodic narrative. Patterns of disequilibrium and equilibrium are therefore inherent in communication. Inversion then acts as the

metaphoric correlate of these forces. It is iconicity which binds together the physiological and cognitive structures which govern prosody.

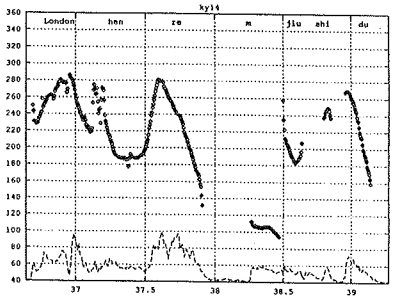


**Figure 1.**

*London hen3 re4 jiu3 shi2 du4*
*[London's very hot, ninety degrees]*

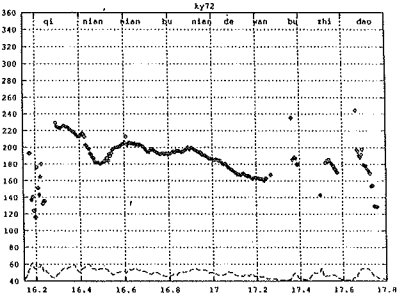


**Figure 2.**

*qi1 nian2 nian4 bu2 nian4 de0 wan2*
*bu4 zhi1 dao4*

**ACKNOWLEDGEMENTS**

This research was conducted at the Phonetics Lab of Stanford University. I am grateful to Bill Poser for his generous support, to Tom Veatch for help with the WAVES program, and to Stanford Linguistics Department for supporting this research.

**REFERENCES**

D. Bolinger (1989), *Intonation and It Uses* (Stanford University Press).

Y.-R. Chao (1968), *A Grammar of Spoken Chinese*. (University of California Press, Berkeley).

I. Fonagy (1971), "Double coding in speech", *Semiotica* 3:189-222.

L.-C. Yang (1992), "A semantic and pragmatic analysis of tone and intonation in Mandarin Chinese", in *ICSLP 92 Proceedings* (University of Alberta Press), pp. 655-658.