

# Harmony and Tension in Mandarin Chinese Prosody: Constraints and Opportunities of Lexical Tones in Discourse Markers

Li-chiung Yang

English Language Center, Department of Foreign Languages & Literature,  
Tunghai University, Taichung, Taiwan

yang\_lc@thu.edu.tw, lyang@sprynet.com

## Abstract

Prosody in tonal languages such as Mandarin provides a fascinating test on the universal character and attributes of prosody in natural language. Prosodic variation is a key element in marking intention, cognitive states, and topic development, and answers to how these communicative goals are accomplished in Mandarin provide enlightening discoveries on the universality of prosodic forces and shapes. This paper analyzes prosody and tonal interaction in the use of two most frequently occurring discourse markers in Chinese, *ranhou* 然後, “then” and *jiushi* 就是 “that is”, each of which plays a critically important role in signaling relationships among topic elements and the mental and interactional states of participants during conversations. The current study investigates how the lexical tones and usage of these markers interact with prosody to bring about the desired communicative goal in spontaneous natural conversation.

**Index Terms:** prosody, universality, lexical tones, discourse markers, Mandarin

## 1. Introduction

Communication in spontaneous conversation is a highly complex process of presentation of topical information, sharing of cognitive and emotional judgments, and on-going interactive signals of adequacy of knowledge state and the desired direction of the conversation. Cognitively, participants are constantly engaging in interpreting and analyzing interactive give-and-take, as well as retrieving relevant information into active memory. Given the complex nature of natural conversation, two resources that are key to how participants successfully build and organize a conversation are lexical-verbal guideposts and prosodic signals.

Discourse markers constitute the key set of lexical-verbal organizational guideposts to idea relationships and topic development [1] [2]. They have been described by [1] as “members of a functional class of verbal and nonverbal devices which provide contextual coordinates for ongoing talk”. The importance of discourse markers arise because they act to provide unit-to-unit coherency through appeals to mental constructs of conventionalized relationships among ideas. Discourse markers include connectives such as *and* and *but*, signals of inference such as *then*, and feedback markers [3]. In Mandarin, there is a similarly rich set of discourse markers including feedback such as *dui* “right”, *na* “well”, *nema* “in that case/therefore”, *jiushi* “that is”, and lexical time markers such as *ranhou* “then/afterwards”. A number of researchers using a discourse approach have studied *ranhou* and *jiu* or *jiushi* [4] [5] [6] [7], but the prosodic characteristics of these markers, and prosody’s role in their

discourse functions has not yet received a systematic integration.

Prosody at first glance appears to act as an independent expressive and organizational resource that accompany the lexical stream of words, but that simultaneously provide a powerful and deeply rooted oral image on a speaker’s cognitive state, emotional reaction, and indications of prominence and importance. Prosody has been seen to arise from such varied aspects of communication as sentence structure, grammatical and topic structure, specific emotions and cognitive states, and pragmatic and interactive signals of speaker conversational intention.

Tone languages such as Mandarin Chinese present a fascinating window on the universal nature of prosodic phenomena. In Mandarin, lexical tones differentiate semantic meaning through variations in pitch shape among four defined tones and a neutral tone. How does Mandarin communicate the emotional states and attitudes, as well as signals of prominence and discourse structure, that are communicated by prosody in other languages, given the seeming restrictions on pitch movement of the lexical tones? Although a full characterization of how Mandarin provides communicative and expressive signals in an environment of lexical tones is only at the beginning stages, the current study forms part of an overall project on discourse, cognition, and prosody, focusing on a key set of conversational linkage points: discourse markers.

## 2. Approach

In the current study, we explore the question of prosody, tones, and the universality of prosodic elements in the specific case of two Mandarin discourse markers, *ranhou* 然後 “afterwards-then” and *jiushi* 就是 “that is”. We use an approach developed in the discourse analysis field, as well as utilize tools and techniques from speech technology and corpus-based linguistics. By using both acoustic and discourse data, we identify patterns of systematic variation in pitch shape in the presence of specific grammatical, cognitive, topic, and interactive influences. In the current study, we explore the effects on the defined tonal shapes under different cognitive and discourse elements.

## 3. Methodology and data

The data analyzed in this study is a subset of data collected under a Taiwan National Science Council funded project on Mandarin discourse markers. The data consist of 10-12 pairs of spontaneous conversations between friends on topics selected spontaneously by the participants. All conversations were transcribed with respect to key elements by student assistants. The conversation analyzed for this paper totaled 60 minutes of continuous conversation between 2 participants.

Speech data were processed using WaveSurfer and ESPS Waves+ software. These conversational data were analyzed by two analysts and annotated for discourse relations, discourse markers, and speaker turns. Instances of the target tokens were extracted from the corpus and analyzed both acoustically and contextually. The acoustic measurements of pitch (f0) and duration of each instance of the two target discourse markers were correlated with the specific characteristics of the discourse and interactive context. In our future work, we intend to include more analysts and annotators so more concrete aspects of meaning differentiation can be derived.

#### 4. Tonal and prosodic characteristics of the discourse markers *ranhou* and *jiushi*

The two discourse markers studied in this paper are key markers of idea unit relationships in Mandarin conversation, but have opposing and distinct usages in discourse. They both introduce new information and point forwards and back to connect topic statements. They occur at critical topic and idea turning points of conversations, and therefore are also elements of floor-negotiation signaling. Mandarin *ranhou* “afterwards-then” is parallel to the English marker *then* in usage, and points to further development in the ensuing conversation. *Ranhou* indicates a moving on from a previous step or idea, either in chronological or discourse time. The marker *jiushi* “it/that is”, on the other hand, points to a new specification or clarification of an immediate presented idea or information, and we can expect a reformulation of the idea or a reference to mutually known information, often with the intent to clarify a point or make more specific an inadequately grasped point.

##### 4.1. Sequential discourse development: the marker *ranhou* “afterwards-then”

The marker *ranhou* has several important discourse functions that justify its very frequent use in spontaneous conversations. Previous research has indicated that *ranhou* acting in its semantic role “afterwards” links units of temporal sequences. Similarly to English “then”, *ranhou* also acts to link idea units and event sequences to signal the next step in a logical development. Because of its role as a sequential link, in spontaneous conversation *ranhou* occurs at points where speakers are actively constructing the ongoing topic flow under conditions of uncertainty with respect to drawing elements into active memory, interpreting signals of comprehension or incomprehension from the hearer [8] [9]. In the process of constructing a coherent flow of ideas, *ranhou* holds a key grammatical position where the possibility of uncertainty and need for immediate resolution of uncertainty is at a high point. Our data indicate that the degree of uncertainty is a key factor in the prosody of *ranhou* and in how the lexical tones of *ranhou* are modified prosodically under conditions of uncertainty.

As generally acknowledged, *ranhou* has a 2-4 lexical tone pattern on its two syllables, a 2<sup>nd</sup> tone rise followed by a 4<sup>th</sup> tone fall. Our data show this tone pattern is influenced both by its temporal or discourse sequential character, by emphasis, and by the degree of uncertainty. To a significant degree, length of utterance is also correlated with uncertainty. In Figure 1 and Figure 2, we show the pitch tracks of instances of *ranhou* for two speakers, M and B, where the duration of *ranhou* is very short. Our analysis indicates that short duration *ranhou*'s occur when the level of mutual understanding of speakers and hearers is at a high level, and

there is little cognitive difficulty in presenting succeeding ideas. Figures 1 and 2 show that for both speakers 2<sup>nd</sup> tone *ran*'s rise is preserved in almost every instance, but there is no significant corresponding tonal fall for 4<sup>th</sup> tone *hou*.

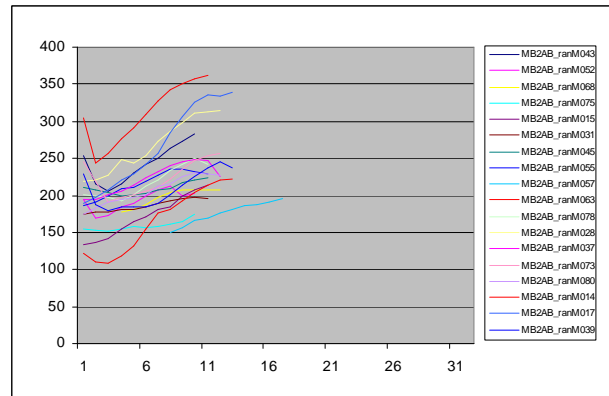


Figure 1: Short-rising pitch pattern for *ranhou* for speaker M

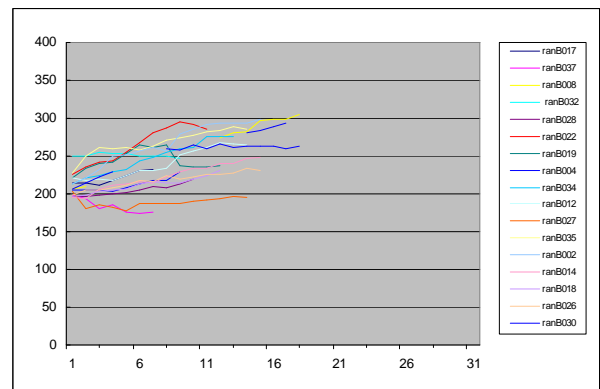


Figure 2: Short-rising pitch pattern for *ranhou* for speaker B

When there is a focus on the temporal sequence nature of the relationship connecting two idea or narrative units, our data indicate that this focus is frequently brought about through both a lengthening of duration and a greater articulation of the full lexical tone values. The arched pitch shapes in Figure 3 show the fuller articulation of the rising – falling lexical tone pattern of *ranhou* for speaker M. In general, our data suggest that fuller articulation of tonal shape is a frequent device to bring attention to and emphasize the lexical meaning of a linguistic token, and such a phenomenon would represent the opportunities for prosody that a tonal language such as Mandarin have that are not present in non-tonal languages. The lexical tones can represent a restriction on prosodic patterns, but they also provide opportunities for prosodic usage by greater-than-normal emphasis on the lexical tonal values. This can be especially powerful when the tonal and prosodic pitch tendencies coincide. When a speaker has strong certainty about a sequential relationship, the prosodic tendency for falling pitch to signal and express definiteness coincides with the 4<sup>th</sup> tone falling tonal value for *hou* in *ranhou*, but is frequently accentuated.

Our contextual analysis indicates that when uncertainty is greater, the duration of *ranhou* is correspondingly lengthened. The lengthened duration provides more time for the speaker to retrieve relevant information into active memory, and to construct the next sequential or logical idea unit in the discourse flow [8] [9] [10] [11]. The lengthened

duration also acts as a floor-holding device. By contrast to the falling pitch *hou* for emphasis and certainty, the longest duration *ranhou*'s were all produced under conditions of speaker uncertainty. Figure 4 shows the rising pitch levels across the extent of *ranhou* for speaker B in such uncertain situations.

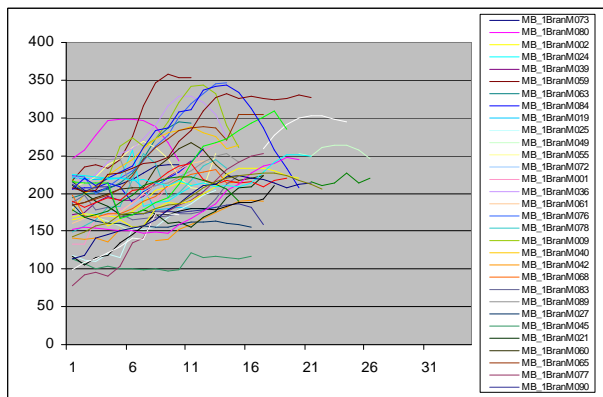


Figure 3: Instances of medium duration, rising and rise-fall pitch patterns for speaker M

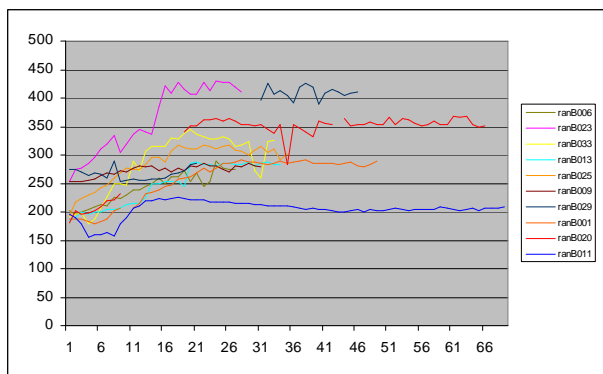


Figure 4: Instances of long rising pitch pattern due to uncertainty and floor negotiation for speaker B

Uncertainty can arise from difficulties in memory retrieval of ideas, relationships, or choice of the correct way of stating an idea, and may also be initiated because the speaker has detected that the hearer has not achieved an adequate comprehension of the point made. In the latter case, speakers may respond not by *pointing forward* to the next idea unit as with *ranhou*, but by *pointing back*, with the marker *jiushi*.

#### 4.2. Specificity and focus: the marker *jiushi* “that is”

By contrast to *ranhou*, which implicitly signals the completion of a prior idea by pointing to the next, the marker *jiushi* explicitly links back to a just prior idea or unit of information by making the prior expression more specific, more definite, and usually more concrete. In our data, *jiushi* arises when there is some inadequacy or incompleteness that needs resolution, and the resolution is commonly phrased in terms of mutually shared knowledge that will help the hearer to understand the point more fully.

Thus, *jiushi* would at first glance appear to occur as a more definite, more explicit, and more certain, expression. Although our data indicate that this is mainly the case, there are also cases where *jiushi* occurs as uncertain or hesitant

when the speaker tries to reformulate a more specific statement, but is temporarily uncertain how to do that.

Figure 5 shows examples of short duration *jiushi* for speaker M. In each instance, falling lexical 4<sup>th</sup> tone for *jiu* becomes flat or slightly rising, and about half the pitch shapes for falling 4<sup>th</sup> tone *shi* are flat. Usage of *jiushi* in these short duration cases is parallel to short duration *ranhou* in some aspects: the speaker does not encounter cognitive difficulties and idea units flow smoothly from one to the next. Comparing the two discourse markers, the rising-falling tone unit, *ranhou*, became primarily rising, and the falling-falling tone pair, *jiushi*, became primarily flatter, in a similar environment. One can infer that in smooth flowing speech areas, without disruptions by uncertainty or specific emotions or emphasis, there is a prosodic tendency for a higher pitch slope for these markers at the linking points of sequential idea units. This is likely to be due to the presence of a continuation function, which both of these markers share. Raising the pitch slope has differential effects because the two markers differ in their lexical tone structure, but the force of prosodic rising under continuation is evidence of the universality of this underlying prosodic feature.

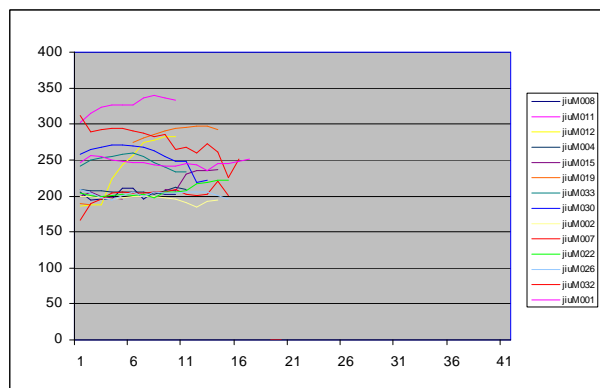


Figure 5: Instances of short *Jiushi* of speaker M, showing 4<sup>th</sup> tones preserved more on *shi* than on *jiu*

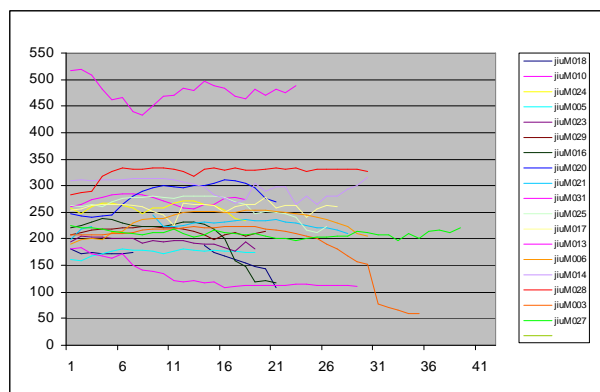


Figure 6: Instances of flat-fall pitch patterns of *jiushi* for speaker M

Markers occurring with longer duration are less likely to act merely as simple continuation markers, and more likely to contain a greater component of pragmatic or semantic content. Figure 6 shows further examples of the pitch movement of *jiushi* for speaker M. The lexical pitch values are much more evident, especially on the latter *shi* syllable. In this section of the dialogue, speaker M is speaking in a much

more expository and definite speaking style, with only several instances of hesitation while searching for words.

Speaker characteristics and a speaker's role at a particular point in a conversation also influence the specific pitch realization of the lexical tones. Figure 7 shows examples extracted from one area of the conversation where speaker B is speaking in an animated fashion and with a high degree of excitement and urgency. The high pitch levels and long duration instances are used primarily as signals of urgency and as a strategy to hold the floor in the threat of interruptions, so that an urgent point can be completed. As shown in Figure 7, lexical 4<sup>th</sup> tone *shi* still exerts a falling-pitch influence over the final portions of the two syllables.

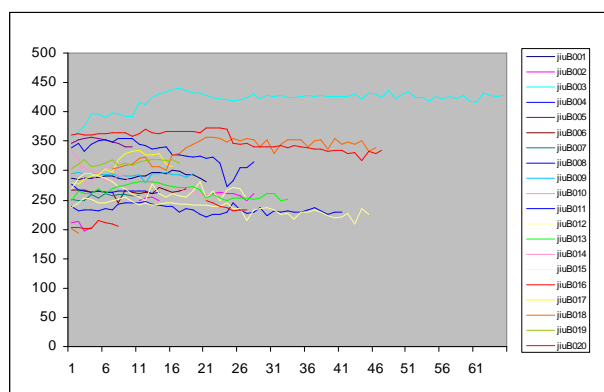


Figure 7: Varied pitch pattern realizations of *jiushi* for speaker B

Because *jiushi* has a reformulating and completing function, and is primarily used from a position of greater relative knowledge on the part of the speaker, *jiushi* more frequently accommodates to the universal effect of definiteness and completion to lower the pitch, by contrast to *ranhou*. The synchrony of lexical pitch and universal prosodic influence in the case of *jiushi* helps to provide the contrasts between the realized pitch patterns of these two contrasting and complimentary markers

## 5. Conclusion

The markers *ranhou* and *jiushi* are two key discourse markers that introduce new information and construct linkages from just presented information to upcoming new information. However, the nature of the information introduced differ in complementary ways. While *ranhou* marks closure of previous information and introduction to a new next step, *jiushi* primarily provides a reformulation or more adequate account of the previous point. This difference accounts for the differing effects of universal prosodic forces on the realized pitch values of the tones.

In this paper we have shown that in a tone language as Mandarin Chinese, tonal patterns are not held rigidly as previously hypothesized, but they fluctuate according to universal prosodic influences, and that because of the differing pragmatic, intentional, cognitive, and interactional uses of these two tokens, the effects on pitch shape variation also differ. The data presented provide evidence of the universality of prosody, even under the constraints of lexical tones, and indicate that the final realized pitch value carries the effects of both tones and prosody. Finally, we also propose that the presence of lexical tones provide unique opportunities for prosodic expression that are not applicable

to non-tonal languages, because of the ability to *utilize* the tone-prosody dichotomy to expressive advantage.

## 6. References

- [1] Schiffrin, D., *Discourse Markers* [Studies in Interactional Sociolinguistics 5], Cambridge: Cambridge University Press, 1987.
- [2] Fischer, K. & Brandt-Pook, H., "Automatically Disambiguating Discourse Particles", Proceedings of Coling/ACL '98 Workshop on Discourse Relations and Discourse Markers, Montreal, Canada, pp. 107-113, 1998.
- [3] Heeman, P.A., and Allen, J. F., "Speech Repairs, Intonational Phrases and Discourse Markers: Modeling Speakers' Utterances in Spoken Dialog", *Computational Linguistics*, Vol. 25-4:1-45, 1999.
- [4] Biq, Yung-O. *The semantics and pragmatics of Cai and Jiu in Mandarin Chinese*. Ph.D. dissertation. Cornell University, 1984.
- [5] Hsieh, Fuhui. "Mandarin Adverbial *Jiu* in Discourse". Proceedings of the 17<sup>th</sup> Pacific Asia Conference on Language, Information, and Computation, 2003.
- [6] Liu, Mei-chun. *Discourse, grammar, and grammaticalization: synchronic & diachronic analyses of Mandarin adverbial markers *Jiu* & *Cai**. Ph.D. dissertation. University of Colorado at Boulder, 1993.
- [7] Su, Lily I-wen. "Rethinking Sequentiality and Coherence: A study based on Chinese connectives". Paper presented at the 7<sup>th</sup> International Cognitive Linguistics Conference, UC-Santa Barbara, 2001.
- [8] Yang, Li-chiung. *Intonational Structures of Mandarin Discourse*. Ph.D. dissertation, Georgetown University, 1995.
- [9] Yang, Li-chiung. "Integrating Prosodic and Contextual Cues in the Interpretation of Discourse Markers". In Kerstin Fischer (ed.), *Approaches to Discourse Particles. Studies in Pragmatics, 1*. Amsterdam: Elsevier. pp. 265-298, 2006.
- [10] Chafe, Wallace. "Cognitive constraints on information flow. In Coherence and Grounding in Discourse", R. Tomlin (ed.). Amsterdam and Philadelphia: John Benjamins, 1987.
- [11] Chafe, Wallace, "Prosodic and Functional Units of Language", in Talking Data: Transcription and Coding in Discourse Research, Jane Edwards and M. Lampert (eds.), Hillsdale, N.J.: Lawrence Erlbaum, 1992.
- [12] Kerstin Fischer. "Frames, constructions and invariant meanings: The functional polysemy of discourse particles". In Approaches to Discourse Particles, number 1, Studies in Pragmatics. Amsterdam: Elsevier, 2006.