PHONOLOGICAL INFLUENCES ON THE REALIZATION OF FINAL LOWERING: EVIDENCE FROM DIALOGUE CHINESE MANDARIN

Wei Lai1, 2, Ya Li2, Xiaoying Xu1, 2, Hao Che2, Shanfeng Liu2, Jianhua Tao2

1 School of Chinese Language and Literature, Beijing Normal University, Beijing, China
2 National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science, Beijing, China
wlai@mail.bnu.edu.cn, {yli,hche,sfliu,jhtao}@nlpr.ia.ac.cn, xuxiaoying2000@bnu.edu.cn

ABSTRACT

Despite the discovery of final lowering effect in widespread language, its origin and realization in different phonological environments still needs exploration. In this article, with a large dialogue corpus, three experiments are conducted to examine how phonological factors (such as prosodic units, sentence stresses and boundary pitch movement) would influence the realization of final lowering in Chinese Mandarin. The results show that: I) The bearing unit of final lowering in Chinese is the last prosodic word in the utterance, regardless of its length, rather than a fixed duration range in a physiological way. II) The position of the sentence stress has an influence on the presence/absence of final lowering. To be specific, final lowering tends to be triggered by sentence stresses on the penultimate and last third prosodic word, and suppressed by sentence stresses prior to the last third prosodic word. III) Final lowering effect would be pushed leftward by sentence stresses and high boundary tones in final positions. This article lends support to the phonological origin of final lowering, and introduces a cross-linguistic framework of prosodic structure to analyze its specific realization under different conditions of stress positions and boundary pitch movements.

Index Terms— Final lowering, Chinese Mandarin, interrogative sentences, dialogue speech, the prosodic structure

1. INTRODUCTION

Final lowering, indicating an additional lowering near the end of the utterance, has been found in a variety of languages with very different prosodic systems, such as English [1-3], Greek [2], Danish [4], Dutch [5], Yoruba [6, 7], Spanish [8], German [9], Japanese [10, 11], Kipare [12] and Chinese [13]. Despite the discovery of final lowering effect in widespread languages, it still remains disputable whether its origin is physiological, phonological or paralinguistic.

Potential physiological triggers of final lowering were first discussed by Liberman and Pierrehumbert [1]. They thought that a drop in the subglottal pressure and relaxation of the laryngeal muscles were necessary articulatory correlates to final lowering. This claim was followed by Herman, Beckman and Honda [14], who considered final lowering as one of the products of an overall “vocal effort” decrease towards the end of an utterance. This decrease would result in a drop in both subglottal pressure and gesture stiffness, and cause a series of chain reaction including segmental lengthening, average amplitude decreasing, voice quality changing and final pitch lowering.

Meanwhile, linguistic explanations from different languages were also continuously put forward. One potential piece of evidence is its close relationship with other prosodic factors. Arvaniti found that final lowering was prone to target the last phrase accent regardless of its relative distance to the end of the utterance [3]. Another support is its connection with declaratives. In Kipare, final lowering was found absent in sentences meant to show incredulity [12]. The third is its potential link to focus encoding, more specifically, the post focus compression [15, 16]. Pierrehumbert and Beckman studied Japanese tone structure and found that final lowering would be triggered by an accentual high tone H!L, but not by a phrase H tone even when it preceded a separate L tone [11]. Similar phenomenon was detected in German [9] and Chinese [17], indicating that final lowering was tied with the extra pitch expansion of its preceding accent. Also involved is the paralinguistic explanation that final lowering undertakes the pragmatic function of signaling discourse finality [18].

There are various methods for detecting final lowering, such as predicting the value of final peaks by data modeling [1, 9], comparing final vs. penultimate peaks [2, 13], comparing F0 drops between successive peaks [19], comparing mean F0 between final vs. penultimate syllables [12], comparing pitch of the same component in final vs. non-final positions [20, 21], etc. Arvaniti compared the effect of the former three methods and suggested the final vs. penultimate peak comparison to be more sensitive and thus more appropriate for final lowering detecting [3], which is
adopted in this article. Another popular topic is the impact of other phonological factors on final lowering, but researches on this issue aim mostly at non-tone languages and focus mainly on the factor of boundary tone combination [9-11, 14, 25].

The present study uses a Chinese dialogue corpus to examine the realization of final lowering in specific phonological environments from the perspective of a tone language. Three experiments are designed respectively to assess how the realization of final lowering is influenced by prosodic units, sentence stresses, and the boundary tone. In the end, this research gives an attempt to discuss about the phonological processing of final lowering within a cross-linguistic framework of prosodic structure (accentual phrase, intermediate phrase and utterance), which makes a contribution to the development of prosodic theories for Chinese Mandarin and to the cross-language comparison of prosodic structures.

2. CORPUS

Our research is based on a large-scale Question & Answer conversation corpus which contains more than 2000 turns of questions and answers. All the conversations are selected from Chinese interview programs and involve a wide range of topics. These conversations were transcribed and then re-read by a trained male speaker in a professional recording studio to make sure that all the F0 values come from the same speaker and are comparable. During the recording process, the speaker was required to maintain a natural speaking style without acting or exaggeration. Four levels of prosodic units (syllable, prosodic word, prosodic phrase and intonation phrase) were annotated and manually checked by the first author.

Different types of interrogatives are contained in this corpus. Some sentences express doubt by lexical means, such as wh- questions (1405 sentences), particle yes-no questions (341 sentences) and v-neg-v questions (184 sentences). The three types of questions convey doubt respectively by wh- words, final neutral-tone particles "吧"/ba/ or "吗"/ma/, and the republication of the predicate-verb. Some express doubt by particular syntactic structures or phrases, such as alternative questions (101 sentences) and tag questions (114 sentences). Some depend exclusively on intonation to express interrogation, such as syntactically unmarked questions (41 sentences). There are also 2179 declarative sentences serving accordingly as responses. In case that the neutral-tone syllables at the end of the utterance would drift down the average pitch values, sentences ending up with the neutral tone were excluded for wh- questions, v-neg-v questions and declaratives, but kept for particle questions which all end up with neutral tone particles.

3. EXPERIMENT AND RESULT I:
BEARING UNIT OF FINAL LOWERING

The first experiment examines whether the basic bearing unit of final lowering is the syllable (SYL) or the prosodic word (PW). Two tasks are contained in this experiment. The first task aims to examine the basic unit of final lowering among different sentence types, i.e., wh- questions, particle questions, v-neg-v questions and declarative sentences. It has already been proved that final lowering exists with these types of sentences in Chinese [13]. The second task aims to examine the same question within a single type of sentence.

3.1. Evidence from Multiple Sentence Types

Sequential peaks of the last six PW and the last six SYL were extracted and plotted for different types of sentences, to check about their final pitch variation. As shown in Fig. 1, the pitch contours on PW peaks are more similar among different types of sentences than those on SYL peaks. It can be observed that final lowering affects intensively the last PW in all sentences, with the final pitch fall 20~30Hz larger than the former ones (except for particle questions whose neutral tone final particles add another significant pitch fall to the end of the utterance). In contrast, pitch variation on SYL peaks shows more randomness. The greatest pitch lowering takes place between the last 2nd & 3rd SYL for wh-questions, and between the last 1st & 2nd and 3rd&4th SYL for v-neg-v and particle questions. For declaratives, the last three syllables are all influenced and show even pitch falls.

![Figure 1. Peak Value Variation on the Last Six PW (left) and that on the Last Six SYL (right) of Wh- Q, Particle Q, V-neg-v Q and Declaratives](image)

Figure 1. Peak Value Variation on the Last Six PW (left) and that on the Last Six SYL (right) of Wh- Q, Particle Q, V-neg-v Q and Declaratives

Further results from post-hoc pairwise comparison following the ANOVA test bring out similar conclusions. Extremely significant pitch differences (Sig.<0.001) take place between the last pair of PW for wh- questions, v-neg-v questions and declaratives, and between the last two pairs of PW as well for particle questions (for the same reason mentioned above). As for SYL comparison, extremely
significance (Sig.<0.001) occurs between last 2\textsuperscript{nd} & 3\textsuperscript{rd}, last 3\textsuperscript{rd} & 4\textsuperscript{th}, last 4\textsuperscript{th} & 5\textsuperscript{th} syllables for wh- questions, between the last 1\textsuperscript{st} & 2\textsuperscript{nd}, last 3\textsuperscript{rd} & 4\textsuperscript{th} syllables for v-neg-v questions and particle questions, and between the last 1\textsuperscript{st} & 2\textsuperscript{nd}, last 2\textsuperscript{nd} & 3\textsuperscript{rd}, last 3\textsuperscript{rd} & 4\textsuperscript{th} syllables for declaratives.

3.2. Evidence from a Single Sentence Type

In this section, declarative sentences are divided into 4 groups based on the length of their last prosodic words, from 1 syllable to 4 syllables. Pitch variation near the end of the sentences of each group is plotted and compared, to see whether the length of the last prosodic word will influence the presence of final lowering.

Fig. 2 shows the pitch variation of PW peaks of declaratives under the conditions that the last PW is made of 1, 2, 3 and 4 SYL. It seems that the last prosodic words tend to undergo a rapid pitch fall of around 20 Hz, regardless of its length, which indicates that final lowering is closely related to the deeper phonological units rather than affect a relatively fixed stretch from the end of the utterance.

Figure 2. Mean Peaks of the Last Six Prosodic Words of Declaratives whose Last Prosodic Words were Respectively Made of 1~4 Syllables

4. EXPERIMENT AND RESULT II: SENTENCE STRESSES AND FINAL LOWERING

This experiment examines the relationship between final lowering and the sentence stress. Pitch differences between the last pair of PW peaks are examined separately when the sentence stress lies in different positions. The 1405 wh-questions are selected as materials to check about this issue. These sentences are relatively shorter and contain less boundary levels (mainly up to prosodic phrase) compared with declarative sentences, which ensures the integrity of intonation over the utterance. More importantly, the accuracy of stress perception is higher in wh- questions than in declaratives, since sentence stresses are prone to target the wh- words which serve as semantic focuses.

Conditions that the sentence stress targets respectively the last PW to the last 6\textsuperscript{th} PW are taken into consideration. For each condition, peaks of the last 6 prosodic words are extracted and compared by post-hoc pairwise comparison after ANOVA. Situations with stresses prior to the last 3rd PW are described in 4.1, those with stresses on the last 3\textsuperscript{rd} and last 2\textsuperscript{nd} PW are described in 4.2., and that with stresses on the last PW is described in Section 5.1.

4.1. Absence of Final Lowering with Sentence Stresses in the Front Part of the Sentence

As shown in Fig. 3, when the sentence stress falls prior to the last 3\textsuperscript{rd} prosodic word, all pitch drops between the last pair of prosodic words are around merely 20 Hz, and no significant final pitch lowering is detected by ANOVA, indicating no additional pitch lowering added to the original declination effect. To be specific, it appears stresses (on the last 4\textsuperscript{th}, last 5\textsuperscript{th}, and last 6\textsuperscript{th} PW) merely bring extra pitch reduction to their subsequent two prosodic words (the last 5\textsuperscript{th} & 4\textsuperscript{th}, last 4\textsuperscript{th} & 3\textsuperscript{rd}, last 3\textsuperscript{rd} & 2\textsuperscript{nd} PW). This reduction seems to directly compress the room of further pitch lowering in final positions.

Figure 3. The Absence of Final Lowering on Sentences with Sentence Stresses Prior to the Third-to-last PW

4.2. Presence of Final Lowering with Sentence Stresses in the Rear Part of the Sentence

Fig. 4 shows that the pitch difference between the last pair of PW reaches its maximum when the sentence stress targets the penultimate PW (about 55 Hz), and reaches the second-largest when the sentence stress targets the last 3\textsuperscript{rd} PW (about 45 Hz). In both conditions extremely significant pitch differences are detected between the last two PW (Sig.<0.001). Meanwhile, pitch of the sentence stress becomes lower when it approaches the end of the utterance.
5. EXPERIMENT AND RESULT III: FINAL LOWERING IN PENULTIMATE POSITIONS

This experiment examines the influence of final sentence stresses and high boundary tones on final lowering. To observe the former condition, wh- questions with sentence stresses on their last PW and alternative questions whose last PW mostly serve as the contract focus are selected as materials. To observe the latter condition, data of syntactically unmarked questions are used, since they rely largely on high boundary tone to convey interrogative mood for lack of syntactic or lexical interrogative markers. The relationship between unmarked questions and the boundary tone enjoys great popularity and has been thoroughly studied in Chinese intonation works via both acoustic and perceptual methods [22-24]. These works basically came to the agreement that the high boundary tone influenced the last syllable of the unmarked questions by heightening its overall pitch and steepening the pitch contour of its lexical tones.

5.1. A Leftward-pushing Effect of Final Lowering Cued by the Final Sentence Stress

The left graph in Fig. 5 shows that a lowering effect occurs to the penultimate PW when the final PW is targeted by the sentence stress in wh- questions. This leftward-pushing effect of final lowering by the final sentence stress also occurs to alternative questions, whose final words are always contrastive and need extra emphasis. As a matter of fact, this phenomenon is not unique to Chinese. Truckenbrodt also found an additional pitch lowering in penultimate positions in utterances ending up with a nuclear accent in German [9]. However, he attributed this lowering to its preceding accents rather than the following nucleus accent.

5.2. A Leftward-pushing Effect of Final Lowering Cued by the High Boundary Tone

Herman found an absence of final lowering in questions to signal incredulity in Kipare, and thought this phenomenon could also be extended to unmarked questions in Mandarin [12]. Nevertheless, Lai suggested that final lowering was supposed to exist in unmarked questions originally, but was offset by the boundary tone effect on the last syllable [13]. Based on Fig. 6, no additional lowering beside declination is detected between the last pair of PW in unmarked questions in the left graph. But when the similar comparison is done to the last pair of SYL, significant pitch lowering is found to target the penultimate SYL, which appears to be pushed leftward by the high boundary tone.
6. DISCUSSION

First of all, conclusions about the bearing unit of final lowering from Experiment 1 provide interesting evidences for settling the dispute between its physiological vs. phonological origins. Previous studies are prone to focus on the issue of “scope” vs “bearing unit”. Conclusions that final lowering targets the last three SYL [12], the last accent [3], and the last PW [13] indicated little difference in the sense of “scope”, but they conveyed totally different information about the “bearing unit”. The former is more physiology related and concerns with the duration of natural pitch decay, while the latter is more of a linguistic question and concerns with prosodic units. (But this does not necessarily mean that it is physiologically irrelevant)

A basic presumption for this issue is: if final lowering is closely related with a “vocal effort” decrease, then it is supposed to target a fixed duration range from the end of the utterance; otherwise, if it tends to target a certain prosodic unit regardless of its length, then the phonological views would make more sense. According to Section 3, final lowering affects merely the last PW in both interrogatives and declaratives no matter how many syllables they are made up of. Chinese is known as a syllable-timed language, in which syllables occur almost at equal intervals. A difference in the number of SYL in a PW might suggest a difference in the real duration of the PW itself. Therefore, it can be concluded that the bearing unit of final lowering is the last prosodic word rather than a stretch at the end of the utterance in the physiological sense. What is more, when pitch variation on a SYL level is examined, conclusions become inconsistent with different types of sentences, which suggests that the bearing unit of final lowering is the last PW rather than the last several SYL in Chinese.

Another finding is the strong relationship between final lowering and positions of sentence stresses. This relationship might be attributed to post focus compression (PFC) proposed by Xu [15, 16], which indicates that the focus can be encoded by a pitch compression of its subsequent components. In this sense, pitch expansion of the sentence stress and compression of its surroundings can be treated as two sides of a coin, tied together to signal prominence on the

6. DISCUSSION

First of all, conclusions about the bearing unit of final lowering from Experiment 1 provide interesting evidences for settling the dispute between its physiological vs. phonological origins. Previous studies are prone to focus on the issue of “scope” vs “bearing unit”. Conclusions that final lowering targets the last three SYL [12], the last accent [3], and the last PW [13] indicated little difference in the sense of “scope”, but they conveyed totally different information about the “bearing unit”. The former is more physiology related and concerns with the duration of natural pitch decay, while the latter is more of a linguistic question and concerns with prosodic units. (But this does not necessarily mean that it is physiologically irrelevant)

A basic presumption for this issue is: if final lowering is closely related with a “vocal effort” decrease, then it is supposed to target a fixed duration range from the end of the utterance; otherwise, if it tends to target a certain prosodic unit regardless of its length, then the phonological views would make more sense. According to Section 3, final lowering affects merely the last PW in both interrogatives and declaratives no matter how many syllables they are made up of. Chinese is known as a syllable-timed language, in which syllables occur almost at equal intervals. A difference in the number of SYL in a PW might suggest a difference in the real duration of the PW itself. Therefore, it can be concluded that the bearing unit of final lowering is the last prosodic word rather than a stretch at the end of the utterance in the physiological sense. What is more, when pitch variation on a SYL level is examined, conclusions become inconsistent with different types of sentences, which suggests that the bearing unit of final lowering is the last PW rather than the last several SYL in Chinese.

Another finding is the strong relationship between final lowering and positions of sentence stresses. This relationship might be attributed to post focus compression (PFC) proposed by Xu [15, 16], which indicates that the focus can be encoded by a pitch compression of its subsequent components. In this sense, pitch expansion of the sentence stress and compression of its surroundings can be treated as two sides of a coin, tied together to signal prominence on the

stressed components. In our case, wh- words in wh-questions often serve as both semantic focuses and sentence stresses, so it is reasonable to presuppose that final lowering can be more easily triggered by preceding sentence stresses due to an additional PFC effect, as mentioned in [9, 17, 20].

Another issue that deserves discussing here, perhaps a more fundamental one, is the status of final lowering in the prosodic structure in Chinese. With reference to the maturely discussed prosodic structure of Japanese and English [11, 25], a prosodic structure contains three levels: utterance (Utt), the highest level concerning with declination, final lowering, and boundary pitch movements (BPM); intermediate phrase (IP), concerning downstep or catathesis in Japanese; and accentual phrase (AP), the lowest level involving phrase-level accents. The three levels are supposed to function independently from the lower to the upper to form the face value of prosody.

This sequence, however, would be problematic with our case if stresses/accents in AP level can exert an influence on final lowering from the upper Utt level. It is risky to arbitrarily classify final lowering to properties of AP, because the sentence stress on the last PW also undergo an extra pitch decrease compared to stresses in other positions, indicating that final lowering is able to affect sentence stresses from an upper level. To address this problem, one must notice that in our case only one sentence stress is annotated for each sentence, which makes it the nucleus rather than accents/stresses in the usual sense. The nucleus, also named sentence/primary/rhematic accent/stresses, is the most prominent stress with a tendency to occur near the end of the utterance [23]. Despite its involvement with AP, it can also be treated as a property of the Utt, for there is only one nucleus per sentence and for its preference for the final positions of an utterance.

This being the case, the influence of sentence stresses on final lowering should be considered to occur thoroughly on the Utt level rather than from lower AP to upper Utt. Also, in this sense the phenomenon that nucleus in front positions fail to trigger final lowering can be explained to some degree. Taking the above analyses into account, the manipulation of final lowering under specific phonological environments can be summarized as Fig. 8.

Figure 8. Phonological Influences on the Realization of Final Lowering in Chinese Mandarin

87
7. CONCLUSIONS

A corpus-based analysis of Chinese dialogue brings about important findings of the phonological control of final lowering. The main results presented here are:

a) The bearing unit of final lowering in Chinese is the last PW, regardless of its length and the sentence type, which lends support to the phonological nature of final lowering.

b) As for the nucleus, final lowering would be triggered if it approaches to the end of the sentence, pushed leftward if it lands on the last PW, and suppressed if it stays away from final positions.

c) High boundary tone in unmarked Q would suppress final lowering in the level of PW peaks, but in fact this lowering effect is pushed to the penultimate SYL peak within the last PW.

The above conclusions would on one hand supply the previous prosodic structure theory with cross-linguistic evidence, and on the other hand contribute to figure out the composition and operation within the prosodic structure of Chinese Mandarin.

8. ACKNOWLEDGEMENTS

The work was supported by the Beijing Normal University (11-30, 10-12-02), the Fundamental Research Funds for the Central University (2010105565004GK), the State Language Commission 12th Five-Year language research programme (YB125-41), the National Natural Science Foundation of China (NSFC) (No.61273288, No.61233009, No.61203258, No.61305003, No. 61332017, and No.61375027), and partly supported by the Major Program for the National Social Science Fund of China (13&ZD189).

9. REFERENCES


