The Role of Prosodic Features in the Perception and Production of Chinese Adverb-Orientation Ambiguous Sentences

Authors: He Jiarui, Zhang Ting

Date: 2025-07-01T00:00:00+00:00

Abstract

Through perception and acoustic experiments, this study investigates the production and perception of adverb-oriented ambiguous sentences by native Chinese speakers and Chinese learners whose native language is non-tonal. Prosodic features play a crucial role in the perception and production of Chinese adverb-oriented ambiguous sentences, specifically manifested as a close correlation between the prosodic features of such sentences and sentence information. During perception, native Chinese speakers can identify changes in prosodic patterns under different sentence meanings and interpret sentence information through prosodic features; during production, they can utilize prosodic features to express different meanings of ambiguous sentences. Chinese learners whose native language is non-tonal have developed an ability to identify different prosodic patterns of adverb-oriented ambiguous sentences that approaches the level of native Chinese speakers, and can to some extent rely on prosodic features to interpret sentence information, but overall find it difficult to perform reasonable disambiguation of adverb-oriented ambiguous sentences through prosodic features.

Full Text

Preamble

Title: The Role of Prosodic Features in the Perception and Production of Adverbially Ambiguous Sentences in Mandarin Chinese

Authors: He Jiarui[123], Zhang Ting[1]
[1] College of Chinese Language and Literature, Nanjing Normal University, Nanjing 210097, China
[2] Department of Chinese Language and Literature, Fudan University, Shanghai 200433, China

[3] International Cultural Exchange School, Fudan University, Shanghai 200433, China

Abstract: This study investigates the production and perception of adverbial ambiguity in Mandarin Chinese among native speakers and learners whose first language lacks lexical tone. Through both perception and acoustic experiments, we examine how prosodic features contribute to the disambiguation of adverbially ambiguous sentences. The findings demonstrate that prosody plays a crucial role in both the perception and production of such sentences, showing a strong association between prosodic cues and sentence-level meaning. In perception, native speakers can identify prosodic variations corresponding to different interpretations and accurately infer sentence meaning based on prosody. In production, they effectively use prosodic cues to convey intended interpretations. While L2 learners approximate native-like sensitivity to prosodic differences and can partially rely on prosodic information for interpretation, they still face challenges in using prosody effectively for disambiguation.

Keywords: prosodic features, ambiguous sentences, perception, production

Introduction

Prosodic features, also known as "suprasegmental features," constitute a phonological structure of language that encompasses variations in acoustic parameters such as pitch and duration beyond segmental quality (Luo Changpei, 2002; Ye Feisheng & Xu Tongqiang, 2011). Ambiguity refers to the phenomenon where a linguistic unit can be interpreted in two or more ways, with common disambiguation methods including phonetic differences, lexical substitution, and syntactic restructuring (Shao Jingmin, 2019). Mandarin Chinese exhibits various types of ambiguity, among which ambiguity arising from different semantic orientations of constituents within a sentence is particularly distinctive. Such sentences share identical structural relationships and hierarchical organization, relying solely on phonetic differences—that is, prosodic features—for disambiguation. The "adverbial orientation" ambiguity discussed in this paper refers to situations where the semantic connection between an adverb and other constituents in a sentence is multidirectional, resulting in multiple semantic interpretations. Shao Jingmin (2019) illustrates this phenomenon using the example "Our school only sent twenty-odd people" : when jiu (only) points forward to "our school," the sentence implies that many people went; when it points backward to "twenty-odd people," it suggests that few people went. Similar adverbs include ye (also), zui (most), dou (all), and cai (only) (Ma Zhen & Lu Jianming, 1988; Zhu Dexi, 1982; Wang Huan, 1999; Zhou Shoujin, 2004; Jin Lixin, 2015).

Previous research has long recognized the interactive relationship between prosodic features and sentence information in adverbially ambiguous sentences. Lü Shuxiang (1999) noted that dou can express generalization, but its scope may include multiple elements, and in certain contexts, it may emphasize one particular element, with stress used to differentiate meanings. Lan Binhan

(1988) emphasized the important role of "logical stress" in disambiguating dou sentences. Chen Xiaohe (1994) described the prosodic features of cai sentences, noting that when cai points forward, sentence stress falls before the adverb, and when it points backward, stress falls after the adverb. Some scholars have employed experimental phonetics to describe acoustic parameters in detail, finding that in dou ambiguous sentences, the adverb' s target exhibits increased duration and fundamental frequency (F0); in ye ambiguous sentences, both forward-pointing and backward-pointing constituents show lengthened duration and expanded pitch range (Yang Yiming, 2000; Wang Ying, 2005); and in cai ambiguous sentences, target constituents demonstrate elevated F0 and duration (Xu Yizhong & Yang Yiming, 2010). Wang Ying (2005) and Huang Caiyu (2013) found through perception experiments that different prosodic patterns of dou and ye ambiguous sentences can be largely understood by information recipients.

Existing research has demonstrated that native Mandarin speakers can effectively utilize prosodic means to disambiguate adverbially ambiguous sentences in spoken production. However, studies on prosodic features in L2 learners' comprehension and production of Chinese ambiguous sentences remain scarce. Chinese learners tend to use only duration changes, pitch variations, or boundary pauses to eliminate ambiguity (Zhou Fengling et al., 2019), yet they frequently exhibit problems such as insufficient duration, missing pitch changes, or incorrect pausing (Liu Min, 2022). Their overall pitch range is also narrower (Liu Shanshan, 2018; Luo Li, 2022). Some learners can interpret syntactically ambiguous dou sentences using prosodic information, but this ability correlates with language proficiency and demonstrates instability—specifically, even learners with extensive Chinese study and high proficiency still struggle to disambiguate dou sentences through prosody at the perceptual level. This is closely related to L1 phonological transfer. As a tonal language, Mandarin' s four tones collectively constitute the vertical dimension of intonation range while also influencing its horizontal contour (Wen Baoying et al., 2022). Mandarin encodes focus through continuous prosodic changes, with focus realized through coordinated variations in duration, pitch, and pitch range. Non-tonal L1 speakers often focus more on pitch changes, resulting in incomplete mastery of Mandarin prosody, including narrow pitch range, incorrect contour patterns, and syllable duration patterns (Li Baogui & Zhou Tiantian, 2020).

Despite extensive research on the prosodic features of ambiguous structures in Chinese linguistics, several limitations exist in the field of international Chinese language education. First, few studies examine L2 learners' perception and production of adverbially ambiguous sentences. Currently, only Yao Qian (2011) has investigated how learners from different L1 backgrounds interpret dou sentences using prosodic features. No relevant research has addressed ye, cai, or zui sentences. Second, existing studies lack a comprehensive perspective integrating perception and production. Research has primarily focused on native speakers' prosodic features when producing adverbially ambiguous sentences, with less attention to how both native speakers and learners interpret such

sentences through prosody. Perception and production represent two crucial aspects of language acquisition with close interconnections. By comprehensively examining learners at different proficiency levels across both dimensions, we can gain deeper insights into their actual language development trajectories and provide valuable guidance for teaching strategies, textbook design, and resource development.

Based on these considerations, this study selects four typical types of adverbially ambiguous sentences—ye, zui, dou, and cai sentences—and employs both perception and acoustic experiments to investigate the role of prosodic features in how native speakers and Chinese learners perceive and produce different meanings in these sentences. We address the following research questions:

1. How do Chinese learners perform when perceiving and producing different meanings of adverbially ambiguous sentences through prosodic features? Are their perceptual and productive abilities consistent? What factors underlie their performance?

1. Perception Experiment

1.1 Materials

Ye, zui, and dou are Level 1 vocabulary items in the *International Chinese Proficiency Standards*, while cai is a Level 2 item. In the grammatical proficiency framework, ye, zui, and dou belong to Level 1 grammar points, and cai belongs to Level 2. For international students, these items present relatively low difficulty and constitute linguistic elements that beginners should master proficiently.

Based on the different orientations of the four adverbs dou, ye, cai, and zui, we created four target sentences with adverbial orientation ambiguity and designed two distinct contexts for each to disambiguate meaning, yielding eight sentences total. To facilitate comprehension for Chinese learners, all lexical items in the target sentences and disambiguating contexts were selected from Level 3 or below in the *Standards*. We controlled for syntactic structure by ensuring each target sentence used a subject-predicate construction with monosyllabic constituents to minimize variable diversity. To ensure material quality, we first invited five undergraduate and graduate students majoring in linguistics to evaluate the ambiguity and naturalness of the materials. Based on their feedback, we revised the materials and finalized the four target sentence sets shown in Table 1 .

Table 1 Target sentences for the adverbially ambiguous sentence production experiment

Target Sentence	Adverbially Ambiguous Sentence
dou sentence	ta dou qu le (He all went)
ye sentence	Zhang laoshi ye jiao Hanyu (Teacher Zhang also teaches Chinese)

Target Sentence	Adverbially Ambiguous Sentence
cai sentence	san ge ren cai chi le yi kuai dangao (Only three people ate one piece of cake)
zui sentence	Anna zui xihuan changge (Anna most likes singing)

Additionally, to ensure naturalness in actual speech flow, we designed disambiguating contexts for more complex ambiguous sentences as two-round dialogues, each containing a question and answer. The first round introduces the topic, while the answer in the second round contains the target sentence. Using "ta dou qu le" as an example, Table 2 presents the dialogue scripts for this ambiguous sentence in different contexts, with scene and role prompts in parentheses.

Table 2 Example sentence "ta dou qu le" matched with different contexts

Forward-pointing context	Backward-pointing context
(Friends chatting)	(Friends chatting)
A: Did Xiao Zhang go to the meeting?	A: Where' s Xiao Zhang?
B: Yes.	B: Xiao Zhang went to eat. What do you need him for?
A: If Xiao Zhang went, I won' t go.	A: I need him urgently.
B: ta dou qu le, can' t you go?	B: ta dou qu le. Wait a bit.

When the adverb dou points to ta (he), ta becomes the discourse center requiring emphasis, meaning "Even Xiao Zhang went, can' t you go?" When dou points to qu le (went), qu le becomes the discourse center, indicating that he has already gone.

We also created four additional sets of eight filler sentences with similar length and structure to the target sentences but without ambiguity, each with corresponding contexts to prevent participants from discovering the experimental purpose. Two undergraduates with Mandarin Proficiency Test Level 2A certification recorded the dialogues in a professional studio, requiring natural conversational speech. We used Audition to extract the target sentences for the perception experiment. Target and filler sentences were presented in random order.

1.2 Procedure

The perception experiment aimed to examine whether L2 learners could identify prosodic features of adverbially ambiguous sentences in different contexts. The experiment comprised two tasks: discrimination and comprehension. The discrimination task assessed whether participants could perceive prosodic feature changes in ambiguous sentences across contexts, while the comprehension task

required participants to match sentences to contexts, investigating whether they could interpret the specific meanings carried by different prosodic patterns.

The experimental design and procedure were as follows: We used PsychoPy 21.2.3 to program the experiment. In the discrimination task, a fixation cue "+" appeared at the screen center for 500ms, followed by a 200ms blank screen. The discrimination task then presented the question: "Are the pronunciations of the audio sentences the same?" Participants could click the audio button to play the sentences repeatedly and press "1" if they were identical or "0" if different, then press Space to continue. We presented four sets of eight target sentences, with each set containing two audio recordings of the same target sentence in different contexts (i.e., with different prosodic features), plus four sets of eight filler sentences where the same audio was played twice. The audio speaker used standard Mandarin with clear articulation and no obvious accent. The order of target sentence sets was randomized.

In the comprehension task, participants listened to audio and judged which of the provided dialogue scenarios the sentence would occur in. Participants completed 16 multiple-choice questions total: eight for target sentences and eight for fillers. Questions for target sentences had one correct answer, while filler questions had all options correct. Dialogue scenarios were familiarized beforehand. The entire experiment lasted approximately 15 minutes.

1.3 Participants

Participants were divided into two groups: native Mandarin speakers and Chinese learners. The native speaker group comprised eight Mandarin-speaking undergraduates majoring in Chinese at Nanjing Normal University, aged 18-20, with no obvious dialect accent and normal language, hearing, and vision abilities. To increase experimental reliability, we additionally recruited 12 native speakers anonymously. For the learner group, we randomly selected 13 Chinese learners from international students at Nanjing Normal University whose native languages were non-tonal, aged 17-24, all having passed HSK Level 3. To determine participants' proficiency levels, we administered an additional language test consisting of a cloze passage using Feng Liping' s (2020) rapid L2 proficiency test. Cloze tests are considered effective measures of learner proficiency (Zhao Yang, 2009) and are commonly used in SLA empirical research (Liu Ying, 2020; Hong Wei & Liu Xiaodi, 2024; He Muxuan et al., 2024). Testing revealed no significant differences among participants, confirming comparable Chinese proficiency levels.

1.4 Results

1.4.1 Discrimination Task Figure 1 [Figure 1: see original paper] shows the accuracy rates of native speakers and learners in identifying prosodic patterns of the four adverbially ambiguous sentence types (dou, ye, zui, and cai sentences) across different semantic interpretations.

Figure 1 Accuracy rates in the discrimination task

The figure clearly shows that both native speakers and learners achieved high accuracy in identifying prosodic patterns for dou, ye, and zui sentences, with cai sentences showing relatively lower accuracy. Notably, both groups achieved 100% accuracy for ye and zui sentences, and over 90% for dou sentences. This indicates that the 20 native speakers and 13 learners could generally perceive prosodic pattern changes in ye, dou, and zui sentences. However, accuracy decreased for both groups when identifying prosodic patterns in cai sentences, with learners showing a particularly marked decline.

1.4.2 Comprehension Task

1.4.2.1 Reaction Time Analysis Table 3 presents the average time required for native and learner groups to comprehend the meanings of dou, ye, zui, and cai sentences under different prosodic patterns, along with between-group comparisons.

Table 3 Reaction time statistics (Note: "1" indicates forward-pointing meaning, "2" indicates backward-pointing meaning)

Sentence Type	Native Group Mean	Learner Group Mean	Significance
dou1	6.05s	15.51s	p<.01**
dou2	8.22s	12.58s	p<.01**
ye1	6.14s	17.03s	p<.05*
ye2	7.29s	15.28s	p<.01**
zui1	4.95s	12.99s	p<.01**
zui2	5.93s	12.17s	p<.01**
cai1	14.04s	24.90s	p=.061†
cai2	13.28s	18.75s	p<.01**

The table clearly shows significant differences in reaction times between the two groups across all sentence types. Learners required approximately 1.5 to 3 times longer than native speakers for semantic processing. For dou sentences, native speakers' average reaction times were 6.05s (dou1) and 8.22s (dou2), while learners' times were 15.51s and 12.58s, respectively, with significant differences (p<.01). For ye sentences, native speakers averaged 6.14s (ye1) and 7.29s (ye2), compared to learners' 17.03s and 15.28s (p<.05 and p<.01). For zui sentences, native speakers averaged 4.95s (zui1) and 5.93s (zui2), versus learners' 12.99s and 12.17s (p<.01). For cai sentences, native speakers averaged 14.04s (cai1) and 13.28s (cai2), while learners averaged 24.90s and 18.75s, with the difference for cai1 approaching significance (p=.061). Overall, native speakers processed all sentence types significantly faster than learners, indicating more efficient

grammatical and prosodic processing. Learners showed clear processing difficulties, requiring substantially more time to understand these sentences, reflecting the challenges of processing complex syntactic structures in L2 acquisition.

1.4.2.2 Accuracy Analysis Table 4 shows accuracy rates for native and learner groups in identifying different meanings of the four sentence types under varying prosodic patterns.

Table 4 Accuracy rates by sentence type (Note: "1" indicates forward-pointing meaning, "2" indicates backward-pointing meaning)

Sentence Type	Native Group	Learner Group	Significance
dou1	0.90	0.69	p=.18
dou2	0.95	0.46	p<.01**
ye1	0.95	0.85	p>.3
ye2	0.90	0.77	p>.3
zui1	0.95	0.84	p>.3
zui2	0.85	0.77	p>.3
cai1	0.90	0.54	p<.05*
cai2	0.60	0.46	p=.493

Native speakers outperformed learners across all sentence types, though significance varied. For dou sentences, native speakers achieved 90% (dou1) and 95% (dou2) accuracy, while learners achieved 69% and 46%, respectively. The difference for dou1 was non-significant (p=.18), but dou2 showed a significant difference (p<.01), indicating learners' substantial difficulty with this pattern. For ye sentences, native speakers achieved 95% (ye1) and 90% (ye2) accuracy, versus learners' 85% and 77%, with no significant differences (p>.3). For zui sentences, native speakers achieved 95% (zui1) and 85% (zui2) accuracy, compared to learners' 84% and 77%, also without significant differences (p>.3). For cai sentences, native speakers achieved 90% (cai1) and 60% (cai2) accuracy, while learners achieved 54% and 46%, with a significant difference for cai1 (p<.05) but not for cai2 (p=.493). In summary, native speakers showed higher accuracy across most sentence types, with significant differences particularly evident for dou2 and cai1, suggesting learners' continued challenges with complex sentence patterns.

2. Production Experiment

The production experiment used the same target sentences and contexts as the perception experiment. Building on the perception experiment, we added four Chinese learners from Nanjing Normal University as speakers, totaling 12 learner speakers whose native languages were non-tonal, all with HSK Level 3 or above and language test scores between 25-35.

Before recording, we explained the materials to participants to ensure familiarity and full comprehension of both meanings of each target sentence. After participants understood the pronunciation and meaning, they recorded dialogues in the phonetics laboratory at Nanjing Normal University' s Suiyuan Campus. Recording software was Adobe Audition 1.5, monophonic, with a 44.1kHz sampling rate, saved as .wav format. Equipment included an HP desktop computer, Neumann U87Ai microphone, and RME Fireface 800 audio interface. Participants unable to come to the laboratory recorded at the Xianlin Campus Arts Center using a Lenovo ThinkBook computer and Logitech Blue Yeti microphone.

The experimenter and participant read scripted dialogues at normal speech rate with natural intonation. The experimenter always read Part A, and the participant always read Part B. The experimenter monitored recordings throughout and required re-recording if the participant made errors or showed obvious disfluency. While minor spontaneous variations in non-target sentences were permitted, target sentences had to match the script exactly.

2.3 Data Collection and Processing

After initial processing of the recorded speech, we used Adobe Audition 1.5 to remove contextual prompts, retaining only target sentences for acoustic analysis.

Given our research needs, we used Praat to annotate speech files at three tiers: (1) Character tier (HZ), marking Chinese characters of the ambiguous structure; (2) Pinyin tier (PY), marking pinyin and tones—since participants produced Mandarin, tones were coded as 1 (yinping), 2 (yangping), 3 (shangsheng), 4 (qusheng), and 5 (neutral tone); (3) Syllable tier (SY), marking syllable boundaries.

We completed segmentation and annotation of target sentences from 25 speakers (8 native, 17 L2 learners), yielding 25$×2×$4 target sentences total. All subsequent prosodic analyses were based on this dataset. We analyzed two prosodic features: fundamental frequency (F0) and duration. All annotated target structure files (including .wav and .TextGrid files) were placed in a single folder. We used Praat scripts to generate F0 files for all extracted target sentences and manually removed extreme outliers deviating substantially from the mean. Since F0 contours were not perfectly time-aligned, we extracted 10 equidistant F0 values within each character of target sentences to enable direct comparison of F0 contours. These values were copied and saved in Excel for processing.

F0 values record actual pitch as produced by individual speakers, which is affected by personal factors. To remove individual variation, we converted F0 values to semitone units using the formula:

$$f(st) = 12 \times \log_2 \left(\frac{f_0}{f_{ref}} \right)$$

where $f(st)$ represents F0 in semitones, f_0 represents F0 in Hertz, and f_{ref} is 50 Hz.

Following the same procedure as F0 extraction, we used Praat scripts to extract syllable durations from the pinyin tier, calculating average duration for each character across participant groups in seconds (s). Converted F0 values, durations, and target sentences were then compiled in Excel for statistical analysis.

2.4 Data Analysis

2.4.1 F0 Features Figures 2-5 [FIGURE:2-5] and Tables 5-8 [TABLE:5-8] present average F0 contours and statistical results for native and L2 speakers reading the four adverbially ambiguous sentence types. In these figures and tables, "a" indicates forward-pointing features and "b" indicates backward-pointing features.

Figure 2 F0 contour for "ta dou qu le"
Figure 3 [Figure 3: see original paper] F0 contour for "Zhang laoshi ye jiao Hanyu"
Figure 4 [Figure 4: see original paper] F0 contour for "Anna zui xihuan changge"
Figure 5 [Figure 5: see original paper] F0 contour for "san ge ren cai chi le yi kuai dangao"

Table 5 Paired-sample t-test results for F0 in "ta dou qu le"

Comparison	Mean Difference	SD	t	p
Forward ta - Backward ta	-4.16	-	-	p<.01**
Forward dou - Backward dou	-4.14	-	-	p<.01**
Forward qu - Backward qu	-0.04	-	-	ns

Table 6 Paired-sample t-test results for F0 in "Zhang laoshi ye jiao Hanyu"

Comparison	Mean Difference	SD	t	p
Forward zhang - Backward zhang	-2.18	-	-	p<.001***
Forward lao - Backward lao	-1.54	-	-	p<.01**
Forward shi - Backward shi	-6.70	-	-	p<.05*
Forward ye - Backward ye	-0.63	-	-	p<.01**
Forward han - Backward han	-0.19	-	-	p<.05*
Forward yu - Backward yu	-2.24	-	-	ns
Forward zhang - Backward zhang	-2.32	-	-	ns

Table 7 Paired-sample t-test results for F0 in "Anna zui xihuan changge"

Comparison	Mean Difference	SD	t	p
Forward Anna - Backward Anna	-4.32	-	-	p<.001***
Forward zui - Backward zui	-5.85	-	-	p<.001***
Forward changge - Backward changge	-2.87	-	-	p<.05*

Table 8 Paired-sample t-test results for F0 in "san ge ren cai chi le yi kuai dangao"

Comparison	Mean Difference	SD	t	p
Forward san - Backward san	-0.82	-	-	p<.05*
Forward cai - Backward cai	-1.19	-	-	p<.05*
Forward yi - Backward yi	-0.26	-	-	ns

The figures and tables clearly show that native speakers and learners produced different F0 patterns for ambiguous structures and the adverbs themselves. Learners' average F0 contours were consistently lower than natives' , indicating generally lower pitch. Pitch has been a primary focus in intonation research (Xie Hong & Shi Feng, 2019), with previous studies generally considering F0 elevation an important acoustic manifestation of emphasizing syntactic structure (Lin Tao, 2013).

Paired-sample correlations were significant for both groups (p<.05). Paired t-tests revealed that native speakers showed significant F0 increases in adverb-targeted constituents when producing adverbially ambiguous sentences. In dou sentences, F0 increases for targeted constituents were highly significant (p<.01). When dou pointed forward, the forward constituent ta showed significantly higher F0 than when it pointed backward. Conversely, when dou pointed backward, the backward constituent qu showed significantly higher F0 than in forward-pointing contexts. Although learners exhibited consistent changes—showing F0 increases in targeted constituents—the magnitude was far smaller than for native speakers.

In ye sentences, when the adverb pointed forward to Zhang, native speakers showed highly significant F0 increases for zhang (p<.001), significant increases for lao (p<.01), and significant increases for shi (p<.05). When pointing backward to han, native speakers showed significant F0 elevation for han (p<.05). Compared to forward-pointing contexts, learners better grasped backward-pointing meanings, as evidenced by significant F0 increases for han (p<.05). In zui sentences, when pointing forward, native speakers showed marginally significant F0 increases for Anna (p=.087). When pointing backward, target constituents showed some F0 decrease, though non-significant (p>.05), while the adverb zui itself showed highly significant F0 elevation (p<.001). Learners showed almost no F0 increase for Anna in forward-pointing contexts but significant F0 increases for changge in backward-pointing contexts

(p<.05). In cai sentences, native speakers showed significant F0 increases for both forward constituent san and backward constituent yi (p<.05). Learners' F0 contours for both meanings were nearly identical, failing to show native-like acoustic patterns at semantic orientation points.

2.4.2 Duration Features Figures 6-9 [FIGURE:6-9] and Tables 9-12 [TABLE:9-12] present duration characteristics and statistical results for native and L2 speakers reading the four sentence types.

Figure 6 Duration statistics for "ta dou qu le"
Figure 7 [Figure 7: see original paper] Duration statistics for "Zhang laoshi ye jiao Hanyu"
Figure 8 [Figure 8: see original paper] Duration statistics for "Anna zui xihuan changge"
Figure 9 [Figure 9: see original paper] Duration statistics for "san ge ren cai chi le yi kuai dangao"

Table 9 Paired-sample t-test results for duration in "ta dou qu le"

Comparison	Mean Difference (ms)	SD	t	p
Forward ta - Backward ta	74.14	10.42	-	p<.001***
Forward dou - Backward dou	14.41	-	-	ns
Forward qu - Backward qu	-63.18	-6.17	-	p<.001***

Table 10 Paired-sample t-test results for duration in "Zhang laoshi ye jiao Hanyu"

Comparison	Mean Difference (ms)	SD	t	p
Forward zhang - Backward zhang	4.41	9.97	-	ns
Forward lao - Backward lao	20.14	0.29	-	ns
Forward shi - Backward shi	27.70	1.28	-	ns
Forward ye - Backward ye	24.66	-24.61	-	ns
Forward han - Backward han	-16.07	-6.35	-	ns
Forward yu - Backward yu	17.91	-1.62	-	ns

Table 11 Paired-sample t-test results for duration in "Anna zui xihuan changge"

Comparison	Mean Difference (ms)	SD	t	p
Forward Anna - Backward Anna	28.22	36.40	-	ns
Forward zui - Backward zui	-3.37	-	-	ns
Forward changge - Backward changge	-	-	-	ns

Table 12 Paired-sample t-test results for duration in "san ge ren cai chi le yi kuai dangao"

Comparison	Mean Difference (ms)	SD	t	p
Forward san - Backward san	29.76	-7.89	-	ns
Forward cai - Backward cai	11.16	25.09	-	p<.05*
Forward yi - Backward yi	-	36.17	-	p<.05*

Duration, as a crucial component of prosody, plays an important role in highlighting semantics. Chao Yuen-Ren (1968) noted that Chinese stress primarily involves expanding pitch range and lengthening duration. Li Aijun (2002) experimentally demonstrated that stressed syllables in Mandarin have longer durations than unstressed syllables. The figures show that in native speakers' forward-pointing productions ("native a"), forward constituents ta, zhang, san, and Anna had longer average durations than in backward-pointing productions ("native b"), while backward constituents qu, han, and changge showed the opposite pattern. Thus, native speakers can disambiguate sentences by lengthening the duration of adverb-targeted constituents to express different meanings. For learners, durations of all constituents were generally longer than for natives, likely reflecting difficulties in lexical and semantic processing requiring more time.

Across the four sentence types, only "ta dou qu le" showed significant duration changes in adverb-targeted constituents for native speakers. When dou pointed forward, the forward constituent ta was nearly twice as long as in backward-pointing contexts. When dou pointed backward, the backward constituent qu was significantly longer than in forward-pointing contexts. Paired t-tests confirmed significant duration increases for targeted constituents (ta and qu) when dou pointed to different elements (p<.001). Learners showed opposite patterns to native speakers in dou sentences.

For ye and zui sentences, native speakers showed consistent duration changes in targeted constituents alongside F0 changes. In ye sentences, when pointing forward to Zhang laoshi, lao showed marginally significant duration increase (p=.065), though the logical stress-bearing zhang showed smaller duration increase than lao. Learners' durations for constituents in both meanings were similar. In zui sentences, when pointing forward to Anna, native speakers showed noticeable duration increase for Anna; when pointing backward to changge, duration increase for changge was not significant. Learners showed trends consistent with native speakers. In cai sentences, native speakers' forward constituents san and yi had longer durations in forward-pointing than backward-pointing contexts, suggesting they cannot use duration lengthening to highlight meaning in cai sentences. Learners again showed patterns opposite to natives.

Regarding the adverbs themselves, zui and dou sentences showed duration changes consistent with F0 changes, with degree adverb zui showing notable

duration increase in backward-pointing contexts. However, ye and cai sentences showed duration changes opposite to F0 changes.

3. Discussion

3.1 Native Speakers' Use of Prosody for Disambiguation

Native Mandarin speakers can rely on prosodic features to disambiguate adverbially ambiguous sentences. At the perceptual level, they can identify prosodic pattern changes across different meanings, with all native participants perceiving changes in dou, ye, and zui sentences and most perceiving changes in cai sentences. Performance was consistent across discrimination and comprehension tasks. The majority of native speakers could understand different meanings conveyed by prosodic variations in dou, zui, and cai sentences and correctly match them to contexts. For the more informationally dense cai sentences, accuracy decreased somewhat for backward-pointing meanings, though many natives could still interpret these meanings through prosody.

At the production level, native speakers could generally disambiguate dou, ye, and cai sentences by highlighting targeted constituents through pitch and duration, consistent with previous findings (Xu Yizhong & Yang Yiming, 2010; Huang Caiyu, 2008; Yao Qian, 2011; Huang Caiyu, 2013). Notably, previous studies used mostly short sentences with 5 or fewer syllables, while our materials contained more syllables and greater processing difficulty. Comparing our materials with previous research reveals that despite higher processing demands, native speakers' ability to use prosody to express semantic focus was not significantly affected, demonstrating that prosodic features are not merely phonetic phenomena but essential tools for effective communication and accurate information transmission in Mandarin, with their function largely unaffected by sentence complexity. When expressing emphasized constituents, native speakers expanded the pitch range of targeted elements while showing substantial post-target pitch range compression and narrowing, consistent with Mandarin focus prosody (Shen Jiong, 1985; Xu, 1999; Wang Bei, 2002; Cao Wen, 2010; Wang Yunjia, 2015).

Previous research did not examine acoustic changes in constituents of zui sentences across different orientations. Our results show that zui sentences share similarities with and differences from ye, dou, and cai sentences. Specifically, when zui points backward, the adverb zui itself shows extremely significant F0 elevation, indicating that native speakers rely more on raising zui's F0 to express backward-pointing meaning—that is, emphasizing that among Anna's many hobbies, singing is the favorite. Notably, in ye sentences, besides highly significant F0 elevation for forward constituent zhang, lao and shi also showed noticeable F0 increases, and duration increase for lao even exceeded that for zhang. This confirms that the intonation center in ye sentences differs from typical logical stress (Yang Yiming, 2000). When pointing backward to Hanyu, han showed significant F0 elevation, matching the logical stress pattern emphasizing that

Teacher Zhang teaches Chinese, not other languages.

Overall, native speakers'duration increases for targeted constituents were smaller than F0 increases. Only dou sentences showed significant duration increases in targeted constituents. Although ye and zui sentences showed duration patterns consistent with F0 patterns, these did not reach significance. Cai sentences showed no duration changes in targeted constituents across meanings. This suggests that native speakers rely more on F0 than duration for disambiguation in natural speech.

3.2 Inconsistency Between L2 Learners' Perceptual and Productive Abilities

Learners could identify prosodic pattern differences in adverbially ambiguous sentences through acoustic cues, with accuracy for ye and zui sentences nearly matching native levels. However, this ability showed imbalance: high accuracy for dou, ye, and zui sentences but notably lower accuracy for cai sentences. When facing specific contexts, learners' ability to interpret sentence information through prosodic cues did not match their ability to identify prosodic patterns, remaining inferior to native speakers. This suggests non-native speakers rely more on physical cues than linguistic information when perceiving Mandarin prosodic boundaries (Chen Mo, 2016).

Across the four sentence types, learners showed relatively accurate understanding of the prosody-semantics relationship in ye and zui sentences, with high accuracy rates. Although learners could identify different prosodic patterns in dou sentences, they still struggled to interpret the meanings these patterns carried, consistent with Yao Qian' s (2011) findings. Learners' low accuracy in both identifying prosodic patterns and interpreting meanings for cai sentences may relate to their complexity and the need to process factors such as subjective quantity comparison (Chen Xiaohe, 1994). Notably, learners showed strong ability to perceive ambiguity caused by degree adverb zui, achieving extremely high accuracy in the discrimination task and high accuracy in comprehending the prosody-semantics relationship for forward-pointing zui. This may relate to our material selection: the forward constituent Anna is a transliteration of a common name across learners' L1s, with an and na mapping to two distinct phonemes in learners' native languages, enabling "two-category assimilation." According to the Perceptual Assimilation Model, L2 learners better perceive L2 sounds that undergo two-category assimilation (Chen Shuwen, 2023). For native speakers, Anna is a less familiar name, potentially requiring attentional resources for recognition and reducing focus on semantic features, thereby decreasing perceptual accuracy. This suggests that incorporating learners' L1 elements in teaching may enhance efficiency.

Production results show that native speakers can disambiguate adverbially ambiguous sentences by producing different prosodic features for targeted constituents, with consistent acoustic manifestations of significant F0 elevation and

duration lengthening. In contrast, learners rarely produced native-like acoustic patterns, showing generally flat intonation without clear stress—a common error in L2 stress production (Deng Dan, 2022). Learners demonstrated weak ability to use prosodic features for disambiguation, failing to accurately grasp discourse centers and logical stress. Even when F0 or duration increases occurred in targeted constituents under certain contexts, these were limited to individual elements and far smaller in magnitude than native productions.

3.3 Explaining L2 Learners' Difficulties in Producing Prosodic Features

Correct production of prosodic features for adverbially ambiguous sentences requires simultaneous attention to sentence meaning and prosodic expression of constituents, posing considerable difficulty for learners. Although learners can partially interpret different meanings through prosody, they often neglect prosodic means for clarifying meaning when processing more complex tasks, adopting avoidance strategies and failing to actively use prosody to accurately express different meanings.

The relatively significant F0 increase for the particle le in learners' dou b productions likely reflects negative L1 transfer. Our learners' native languages (English, Russian, etc.) are morphologically rich, and the meaning expressed by dou b ("he has already gone") can be expressed through grammatical categories of tense and aspect—i.e., verbal morphology—in their L1s. The particle le is often considered a rare "tense-aspect marker" in Chinese (Liang Yinfeng, 2023) that shares similarities with verbal morphology in learners' L1s. In English, ye sentences also create ambiguity. Observing two English-native participants in our study revealed that English speakers indeed outperformed non-English speakers in using prosody to disambiguate ye sentences. Goad & White (2004) proposed the Prosodic Transfer Hypothesis (PTH), suggesting L1 prosodic features affect L2 prosody acquisition. Stress realization differs markedly between non-tonal languages and Chinese, creating difficulties for non-tonal L1 learners. Chinese is a tone-stress language where stress realization must be tone-based (Xu Ximing & Shen Jiaxuan, 2016), whereas our participants' L1s are non-tonal languages where stress depends more on lexical stress than tone. Consequently, learners show weak ability to coordinate lexical tone and intonation when producing Chinese stress, unable to effectively adjust pitch range while maintaining tonal patterns. Even when learners could identify discourse centers and consciously raise F0 of targeted constituents, the magnitude of increase was significantly smaller than for native speakers.

4. Conclusion

Through perception and production experiments, this study examined the acquisition of prosodic features in adverbially ambiguous sentences by Chinese learners whose L1s lack lexical tone. Results show that at the perceptual level, learners can basically identify prosodic pattern changes when adverbs point to

different constituents, but their ability to interpret sentence information through prosody remains relatively weak. At the production level, learners generally struggle to use stress effectively for disambiguating adverbially ambiguous sentences. Overall, learners' perceptual abilities exceed their productive abilities, validating Altmann' s (2006) view that good L2 prosodic perception skills do not necessarily guarantee good L2 prosodic production.

Although many studies confirm the important role of prosody in L2 acquisition, prosodic instruction is often neglected in actual curricula, with some even arguing that pronunciation and intonation cannot be taught. Moreover, prosody acquisition is particularly difficult for learners beyond the critical period (Lengeris, 2012). Chinese learners are typically adults who have passed the critical period, making prosodic instruction especially important. Current prosody teaching in international Chinese education remains unsystematic (Chen Mo, 2021), with teachers often adopting ad-hoc approaches that result in low acquisition efficiency. Therefore, international Chinese teachers should emphasize systematic prosody acquisition, consciously supplementing prosodic knowledge and providing training alongside instruction on correct character and word pronunciation. Training should integrate "perception practice" and "production practice" to help learners improve their ability to discriminate sounds and overcome "foreign accent" issues.

References

[1] *International Chinese Proficiency Standards* (GF0025-2021) [M]. Beijing: Beijing Language and Culture University Press, 2021
[2] Luo Changpei, Wang Jun. *Outline of General Phonetics* [M]. Beijing: The Commercial Press, 2002
[3] Lin Tao, Wang Lijia. *Phonetics Course: Revised Edition* [M]. Beijing: Peking University Press, 2013
[4] Shao Jingmin (Ed.). *General Introduction to Modern Chinese* [M]. 3rd ed. Shanghai: Shanghai Education Press, 2016
[5] Ye Feisheng, Xu Tongqiang, Wang Hongjun, et al. *Outline of Linguistics* [M]. Beijing: Peking University Press, 2010
[6] Ma Zhen, Lu Jianming. *Essays on Modern Chinese Function Words* [M]. Beijing: Peking University Press, 2017
[7] Zhu Dexi. *Lectures on Grammar* [M]. Beijing: The Commercial Press, 1982
[8] Lü Shuxiang. *Eight Hundred Words of Modern Chinese* [M]. Beijing: The Commercial Press, 1980
[9] Wang Huan. *New Chinese-English Dictionary of Function Words* [M]. Beijing: Sinolingua, 1999
[10] Chao Yuen-Ren. *A Grammar of Spoken Chinese* [M]. Translated by Lü Shuxiang. Beijing: The Commercial Press, 1979
[11] Cao Wen. *Prosodic Realization of Focus Stress in Mandarin* [M]. Beijing: Beijing Language and Culture University Press, 2010
[12] Li Aijun. Acoustic manifestation of prosodic features in Mandarin conver-

sation [J]. *Studies of the Chinese Language*, 2002, (6)
[13] Shen Kaimu. On "semantic orientation" [J]. *Journal of South China Normal University (Social Science Edition)*, 1996, (1)
[14] Liu Ningsheng, Qian Yulian. Semantic orientation of zui and entailment of zui sentences [J]. *Chinese Language Learning*, 1987, (5)
[15] Chen Xiaohe. Preliminary exploration of subjective quantity—Also on adverbs jiu, cai, dou [J]. *Chinese Teaching in the World*, 1994, (4)
[16] Zhou Shoujin. Semantic information features of "subjective quantity" and semantics of jiu and cai [J]. *Acta Scientiarum Naturalium Universitatis Pekinensis (Philosophy and Social Sciences)*, 2004, (3)
[17] Jin Lixin. Some explanations on jiu and cai [J]. *Language Teaching and Linguistic Studies*, 2015, (6)
[18] Lan Binhan. Semantic features of adverb dou and its restrictive effect on following verbs [J]. *Language Teaching and Linguistic Studies*, 1988, (2)
[19] Huang Caiyu. Experimental phonetic analysis of semantically ambiguous dou sentences [J]. *Language Teaching and Linguistic Studies*, 2013, (5)
[20] Yang Yiming. On ambiguity in ye sentences [J]. *Studies of the Chinese Language*, 2000, (2)
[21] Xu Yizhong, Yang Yiming. Research on ambiguity and related phonetic issues of jiu and cai [J]. *Language Research*, 2010, 30(1)
[22] Zhou Fengling, Wang Jianqin. Korean learners' Mandarin oral rhythm processing in disambiguating ambiguous sentences [J]. *Second Language Learning Research*, 2019(1)
[23] Deng Dan, Zhu Lin. L2 learners' perception and production of Mandarin neutral tone [J]. *Language Teaching and Linguistic Studies*, 2019, (5)
[24] Wen Baoying, Pan Chaochao, Xu Lizheng. Experimental exploration of Portuguese learners' acquisition of Mandarin declarative intonation [J]. *International Chinese Language Education*, 2022, (3)
[25] Li Baogui, Zhou Tiantian. Experimental study on Italian students' acquisition of Mandarin declarative intonation [J]. *Nankai Linguistics*, 2020, (2)
[26] Yao Qian. Experimental study on using prosodic information to interpret dou sentence ambiguity—Comparison between native speakers and L2 learners [J]. *TCSOL Studies*, 2011, (4)
[27] Feng Liping, Feng Hao, Bai Sida, et al. Development and analysis of rapid Chinese L2 proficiency test—Based on equidistant cloze test [J]. *Applied Linguistics*, 2020, (3)
[28] Zhao Yang. Acquisition of Chinese unaccusative and psychological verbs—On superset-subset relations and learnability [J]. *Chinese Teaching in the World*, 2009, 23(1)
[29] Liu Ying. Thai native speakers' acquisition of Chinese exclusive semantics [J]. *Chinese Teaching in the World*, 2020, 34(4)
[30] Hong Wei, Liu Xiaodi. Effect of metaphorical gestures on Chinese L2 abstract vocabulary learning [J]. *Chinese Language Learning*, 2024(3)
[31] He Muxuan, Zheng Lina, Chang Hui. Interface perspective on Korean learners' acquisition of Chinese negative scope [J/OL]. *Chinese Teaching in the World*, 2024, 38(1)

[32] Xie Hong, Shi Feng. On prosodic encoding of speech information structure [J]. *Journal of Tianjin University (Social Sciences)*, 2019, 21(4)
[33] Shen Jiong. Preliminary discussion on Mandarin intonation models [J]. *Linguistic Research*, 1992, (4)
[34] Wang Bei, Yang Yufang, Lü Shinan. Study on F0 patterns of stressed syllables in Mandarin sentences [J]. *Acta Acustica*, 2002(3)
[35] Wang Yunjia, Ding Duoyong, Dong Tuoxiao. Tone implementation under different intonation conditions [J]. *Acta Acustica*, 2015, 40(6)
[36] Chen Mo. Perception of naturalness in L2 Chinese reading pauses [J]. *Language Teaching and Linguistic Studies*, 2016, (3)
[37] Chen Shuwen. Theoretical foundations and new developments in L2 speech production research [J]. *Contemporary Linguistics*, 2023, 25(4)
[38] Deng Dan. Analysis of Mandarin intonation teaching for L2 learners [J]. *International Chinese Language Education*, 2022, 7(3)
[39] Liang Yinfeng. Semantic evolution direction of Chinese tense-aspect markers le and zhe [J]. *Journal of Nanjing Normal University (Literature and Linguistics)*, 2023, (2)
[40] Xu Ximing, Shen Jiaxuan. Phonological differences in stress between English and Chinese [J]. *Foreign Language Teaching and Research*, 2016, 48(5)
[41] Chen Mo. Prosody: A new focus in SLA research [J]. *International Chinese Language Teaching Research*, 2021, (3)
[42] Wang Ying. Preliminary experimental phonetic study on ambiguous ye sentences [D]. Nankai University, 2005
[43] Liu Shanshan. Analysis of pause acquisition by Russian Chinese learners [D]. Southwest Jiaotong University, 2018
[44] Luo Li. Study on Uyghur students' acquisition of prosody-related interfaces in Chinese ambiguous structures [D]. Sichuan International Studies University, 2022
[45] GOAD H, WHITE L. Ultimate attainment of L2 inflection: Effects of L1 prosodic structure. *EUROSLA Yearbook: 4(1)* [M]. EUROSLA Yearbook, 2004
[46] LENGERIS A. Prosody and Second Language Teaching: Lessons from L2 Speech Perception and Production Research [M]. *Pragmatics and Prosody in English Language Teaching*. Dordrecht: Springer Netherlands, 2012
[47] SO C K, BEST C T. Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories [J]. *Studies in Second Language Acquisition*, 2014, 36(2)
[48] XU Y. Effects of tone and focus on the formation and alignment of F0 contours [J]. *Journal of Phonetics*, 1999, 27(1)
[49] ALTMANN H. The perception and production of second language stress: A cross-linguistic experimental study [D]. University of Delaware, 2006

Note: Figure translations are in progress. See original paper for figures.