

A Prosodic Focus Study on the Differentiation of Syntactic Structural Ambiguity from the Perspective of Experimental Phonetics: A Case Study of “The Mother of Three Children”

Xingwen Xu

*School of Liberal Arts, Heilongjiang University, Harbin, China
cynthia_xingwen@163.com*

Abstract. Research on how prosody guides the parsing of structurally ambiguous strings has increasingly pointed to the role of prosodic focus. Focusing on the English noun phrase “the mother of three children,” we test whether focus placement helps listeners and readers disambiguate structure. We conduct two complementary studies: an acoustic production experiment examining how speakers realize focus on the head noun versus the prepositional phrase, and a perception experiment probing whether those patterns are used as cues to interpretation. Across tasks, focus placement modulates ambiguity resolution: productions show systematic differences in prominence and phrasing when focus targets distinct constituents, and listeners are more likely to adopt the interpretation aligned with the realized focus. The effect is reliable but not uniform across tokens and participants, suggesting that focus interacts with other cues available in context. These findings clarify how prosodic focus functions as a structural cue in real-time processing and delineate the conditions under which it is most informative for syntactic disambiguation.

Keywords: syntactic ambiguity, prosodic focus, experimental phonetics

1. Introduction

Work on syntactic ambiguity increasingly points to prosodic cues as guides to structure, yet we still lack tightly coupled evidence that links where focus is placed to how listeners resolve ambiguity in the same materials. Targeting the English noun phrase “the mother of three children,” this study asks whether focus placement—on the head noun vs. the prepositional phrase—shapes interpretation. We combine an acoustic production experiment, which quantifies duration, F0, and intensity patterns associated with focus, with a perception experiment that tests whether those patterns bias listeners toward one parse over the other. By examining the production–perception alignment within a single paradigm, we specify the conditions under which prosodic focus is informative for structural disambiguation and where its influence weakens relative to contextual cues. The results refine current accounts of prosody–syntax interaction and delimit the contribution of focus to ambiguity resolution in English.

2. Literature review

Scholars have primarily adopted three approaches to analyzing ambiguous sentence structures. The first is hierarchical analysis, which offers an intuitive structural representation of sentences, aiding in

the comprehension of complex sentence constructions [1]. The second is transformational analysis, which demonstrates the possible interpretations of ambiguous sentences through structural transformations [2]. The third is semantic-oriented analysis, which focuses on the semantic relationships between sentence components and is particularly useful for analyzing how adverbial elements with clear semantic tendencies influence overall sentence interpretation [3]. Most existing studies on ambiguity have focused on grammatical, semantic, and pragmatic aspects, requiring advanced linguistic analysis skills and a deep understanding of sentence structures. However, these analytical methods are often complex, time-consuming, and lack phonetic-based analyses and descriptions.

In recent years, research on ambiguous sentence structures has deepened, with an increasing number of studies adopting multi-perspective and multi-method approaches, among which experimental phonetics is one prominent method [4]. Within experimental phonetics, considerable attention has been given to investigating how prosodic focus influences the semantic interpretation of ambiguous sentences. As a high-level manifestation of speech dynamics, prosody involves the coordinated variation of various prosodic elements following specific patterns at the acoustic level. Since Yuen Ren Chao laid the foundation for Chinese prosody research, nearly a century of scholarly efforts has driven the field from predominantly descriptive analyses toward practical applications, marking a shift from qualitative to quantitative research [5]. Prosodic cues primarily include duration, fundamental frequency, and intensity [6]. Lehiste's study found that duration is the most reliable cue in resolving syntactic segmentation ambiguities. Focus represents the central meaning of a sentence as expressed by the speaker and highlights the intended emphasis [7]. Halliday was the first to propose that "focus" corresponds to the most prosodically prominent part of a sentence and serves to convey new information [8]. Huang Caiyu applied the aforementioned prosodic cues in a study on the ambiguous "V double + N double" structure, conducting acoustic experiments and statistical analyses [9], thereby confirming the validity of using duration, fundamental frequency, and intensity to interpret ambiguous sentence structures. Wen Baoying and Dong Weiyan analyzed prosodic focus in the "lian" construction "lian NP dou VP" [10], while Hao Cuixue and Lü Ming examined prosodic focus in the ambiguous "V de A" structure [11]. Both studies found that focus leads to an expansion of the pitch range, while also compressing the pitch range of post-focus elements. This compression effect extends to the end of the sentence, whereas its influence on pre-focus pitch is relatively minimal. These experimental approaches not only provide empirical evidence for understanding linguistic ambiguity but also demonstrate the value of experimental phonetics in exploring complex linguistic phenomena.

The application of experimental phonetics in the study of ambiguous sentence structures holds significant value, as it not only reveals the impact of phonetics on sentence interpretation but also advances the understanding of cognitive mechanisms in language processing. In speech, focus manifests as a synchronic prosodic pattern, yet the role of prosodic cues in cognitive processing remains insufficiently understood in linguistic research [10].

This study aims to bridge this gap by investigating the prosodic focus in syntactic ambiguity resolution from an experimental phonetics perspective, using the ambiguous phrase "sān gè háizi de māma" ("the mother of three children") as a case study. This ambiguous structure has two possible interpretations: (1) "one mother has three children" and (2) "three mothers who have children." The selection of this sentence structure is primarily based on its clear syntactic ambiguity and the potential influence of prosodic focus on syntactic parsing. While previous studies have explored the relationship between prosody and syntactic structure, most have focused on individual factors, lacking empirical research that integrates actual phonetic data.

Analyses of structural ambiguity typically fall into three traditions. Hierarchical (constituency-based) accounts model alternative parses directly in the phrase structure and make the sources of

ambiguity explicit, which helps explain how complex configurations license multiple groupings [1]. Transformational accounts derive competing readings through alternative movement or reattachment operations, thereby exposing which transformations are required for each interpretation [2]. Semantically oriented accounts foreground selectional relations and scope, illuminating how adverbials or other modifiers with strong semantic tendencies bias the global reading [3]. While each approach clarifies a different facet of ambiguity, they often presuppose advanced grammatical analysis and can be costly to apply to spontaneous speech; moreover, their predictions are typically evaluated on written representations rather than on phonetic realizations.

Recent work brings experimental phonetics to bear on ambiguity, asking whether listeners exploit phonetic realizations as cues to structure [4]. Within this line, prosodic focus has received particular attention because it redistributes prominence and phrasing in systematic ways. Prosody, as a higher-level organization of speech rhythm and melody, is realized through coordinated changes in duration, fundamental frequency (F0), and intensity [5,6]. Classic studies suggest that duration often serves as a robust cue in segmentation and grouping, though its effectiveness depends on context and co-occurring cues [6]. In information-structure terms, focus marks what is asserted as new or contrastive; it tends to align with the prosodically most prominent constituent and thus has the potential to guide syntactic parsing [7,8].

A series of production-and-perception studies has operationalized these cues in specific ambiguous patterns. For instance, Huang Caiyu examines the ambiguous “V-double + N-double” configuration and shows, via acoustic measurement and statistics, that duration, F0, and intensity differentiate intended readings to a measurable degree [9]. Wen Baoying & Dong Weiyan investigate focus in the “lian NP dou VP” construction, and Hao Cuixue & Lü Ming study the ambiguous “V de A” pattern; both report focus-induced pitch-range expansion on the focused constituent, accompanied by post-focus compression that can extend to the clause-final region, with relatively minor effects on pre-focus pitch [10,11]. Together, these studies demonstrate that prosodic manipulation supplies useful, albeit not uniform, cues to interpretation.

Despite this progress, two gaps remain. First, prior work often isolates one factor at a time (e.g., duration alone or F0 alone), leaving open how multiple focus-related cues combine during real-time disambiguation. Second, fewer studies track production–perception alignment within a single paradigm on the same materials, which is essential for assessing whether the cues speakers produce are in fact those that listeners rely on. Addressing these points, the present study uses the ambiguous noun phrase “sān gè hái zi de mā ma” (“the mother of three children”) to test whether focus placement on competing constituents systematically shapes both the acoustic realization and the preferred interpretation. By linking cue realization to listener behavior, we clarify when prosodic focus is informative for structural disambiguation and where its influence weakens relative to other contextual information.

3. Acoustic experiment design for the ambiguous structure “the mother of three children”

3.1. Speakers

This experiment involved 12 speakers, consisting of 6 males and 6 females, all aged 20 and from northern Mandarin-speaking regions. Their Mandarin proficiency was at Level 2-A or above. All participants were in good health, had normal articulatory organs, and were fully capable of completing the experimental tasks. They actively engaged throughout the experiment.

3.2. Experimental corpus design

We examined the ambiguous noun phrase “sān gè hái zi de mā ma” (“the mother of three children”) in matched carrier sentences that encourage distinct interpretations while keeping non-target factors

constant. To minimize confounds from segmental content or rhythm, the carriers were structurally parallel and contained the same number of syllables. The two sentence frames were:

Group 1: Tā jiùshì sān gè hái zi de mā ma. (“She is the mother of three children.”)

Group 2: Tā men shì sān gè hái zi de mā ma. (“They are mothers of three children.”)

In the base materials, the target phrase comprises three prosodic words—sān gè (“three”), hái zi (“children”), and mā ma (“mother”)—totaling six syllables. Following Shi Feng, a prosodic word is a perceptually salient grouping with tight phonetic cohesion. Building on this notion, the acoustic analysis quantified duration, fundamental frequency (F0), and intensity both per syllable and across the entire phrase to identify focus placement and to assess the phrase’s internal prosodic segmentation. By keeping the carrier frames parallel and varying only the interpretive context tied to the noun phrase, the design isolates how prosodic focus—as realized in the target constituents—contributes to ambiguity resolution.

3.3. Collection of recording samples

In this study, the Praat software was used for recording, with a mono-channel format and a sampling rate of 44,100 Hz. Praat offers functions for recording, playback, and speech synthesis, allowing for the editing and analysis of selected speech data. The final results can be displayed in various types of spectrograms. The study conducted statistical analyses of the values provided by the software for three fundamental phonetic parameters—pitch, intensity, and duration—while also utilizing spectrograms to analyze the ambiguous structure.

The recordings were conducted in a quiet classroom within the School of Literature at Heilongjiang University. Before recording, speakers were given 1 to 2 minutes to familiarize themselves with the corpus. During the recording process, they were instructed to read the sentences naturally and steadily. Each speaker read each experimental sentence three times, with a 5-second interval between repetitions. The pronunciation quality was assessed by the author of this study, who is also a native speaker from a northern Mandarin-speaking region with a Level 1-B Mandarin proficiency. Data with obvious misinterpretations of ambiguous sentences were excluded, resulting in a total of $10 \times 2 \times 3 = 60$ valid recordings.

3.4. Data measurement and statistical visualization

All recordings were analyzed in Praat to obtain F0 (Hz), duration (ms), and intensity (amplitude sum) per syllable. F0 values were first converted to semitones to reduce scale differences across speakers. To accommodate physiological differences, male and female speakers were processed in separate pitch ranges before aggregation. Unless otherwise noted, measurements are reported at the syllable level and then aggregated to word/phrase levels for comparison across constituents.

Pitch fluctuation was computed following [12]:

$$Q_x = \frac{K_x - K_{\min}}{K_{\max} - K_{\min}}$$

where $K_{\max}$ and $K_{\min}$ are the sentence-level extrema of the (semitone-transformed) F0 range, and K_x is the upper–lower F0 bound measured for syllable x within that range. For each gender group, we plotted per-syllable and word/phrase-level fluctuation curves; thin boxes highlight syllable-level percentage ranges, thick boxes highlight word/phrase-level ranges.

Temporal prominence was indexed with a duration ratio (following the “pause-rate” formulation in [13]):

$$D_x = \frac{S_x}{S}$$

where S_x is the duration of syllable x and S is the mean syllable duration within the sentence. For word-level analysis, we derived:

$$D'_x = \frac{S'_x}{S'}$$

where S'_x is the word duration, and $S' = (\text{mean syllable duration in the sentence}) \times (\text{syllables per word})$. This parametrization allows direct comparison of pre-/post-boundary lengthening and within-phrase timing without conflating speech-rate differences across tokens.

To capture intensity-based prominence, we computed the volume ratio [13]:

$$E_x = \frac{G_x}{G_{\text{mean}}}$$

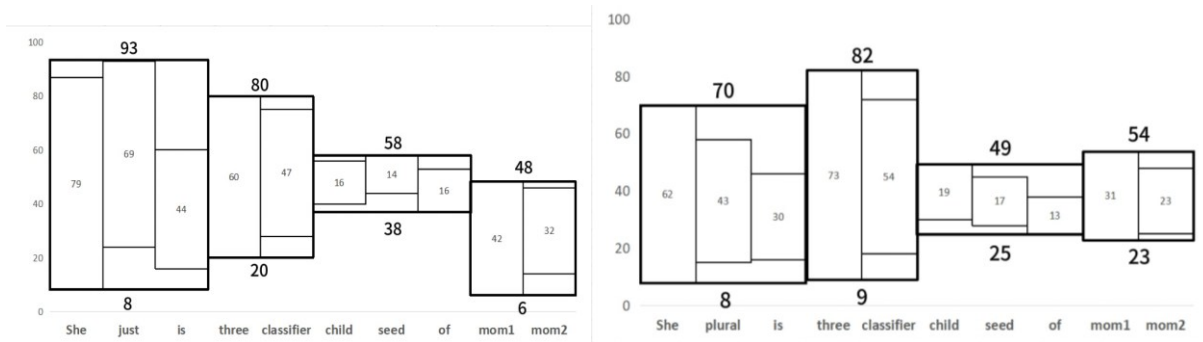
where G_x is the amplitude sum for the target word and G_{mean} is the mean syllable amplitude sum within the sentence. Values greater than 1 indicate increased intensity relative to the sentence baseline. Intensity was extracted syllable by syllable from the waveform and then aggregated to the word/phrase level for comparison.

All descriptive statistics and plots (Q_x , D_x/D'_x , E_x) were produced in Excel, using the same axis scales within each gender group to aid comparability. Summary tables report means and dispersion at syllable and word levels; phrase-level overlays visualize how focus placement aligns with prominence patterns across constituents.

Absolute acoustic values vary across speakers (e.g., by gender and speech rate). Our analyses therefore emphasize within-speaker relative measures (Q_x , D_x/D'_x , E_x) and gender-specific ranges before cross-speaker averaging. While common patterns emerge at the group level, effects are not uniform across all tokens or participants; where relevant, we note deviations and consider possible interactions with contextual cues.

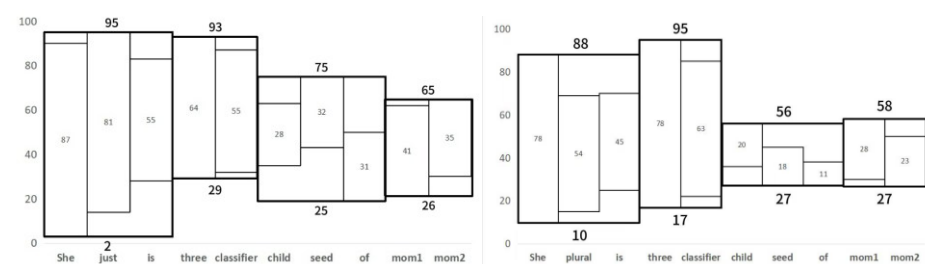
4. Acoustic experimental analysis of the ambiguous structure “the mother of three children”

4.1. Pitch performance under two interpretations



1a. Male Speakers – Ambiguous Sentence 1 (Left)

1b. Male Speakers – Ambiguous Sentence 2 (Right)



1c. Female Speakers – Ambiguous Sentence 1 (Left)
 1d. Female Speakers – Ambiguous Sentence 2 (Right)
 Figure 1. Pitch fluctuation patterns by gender

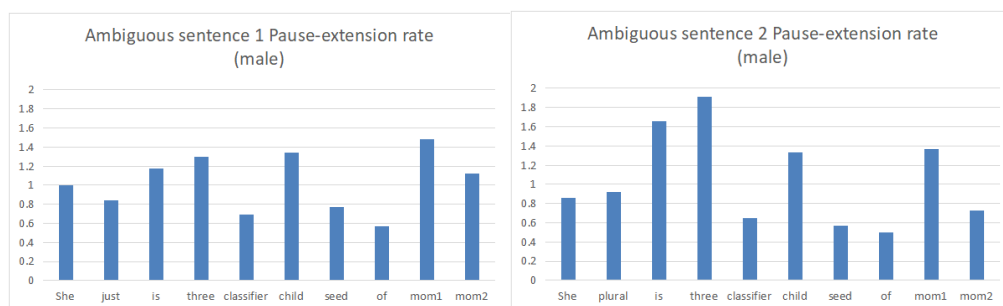
We indexed pitch using F0 fluctuation (Section 3.4) and analyzed male and female speakers separately. Figure 1 summarizes the patterns for the two ambiguous sentences.

In Ambiguous Sentence 2, both male and female groups show greater F0 fluctuation on *sān gè* (“three”) than on *māma* (“mother”) (Fig. 1b, 1d). Within this frame, *sān gè* displays the largest F0 range in the sentence, yielding higher acoustic prominence and a stronger perceptual salience for “three.” By contrast, Ambiguous Sentence 1 exhibits the reverse distribution (Fig. 1a, 1c): *māma* shows a larger fluctuation than *sān gè*, making “mother” relatively more prominent. Thus, across genders, the location of maximal F0 fluctuation tracks the intended focus in each sentence type, aligning with their respective interpretations.

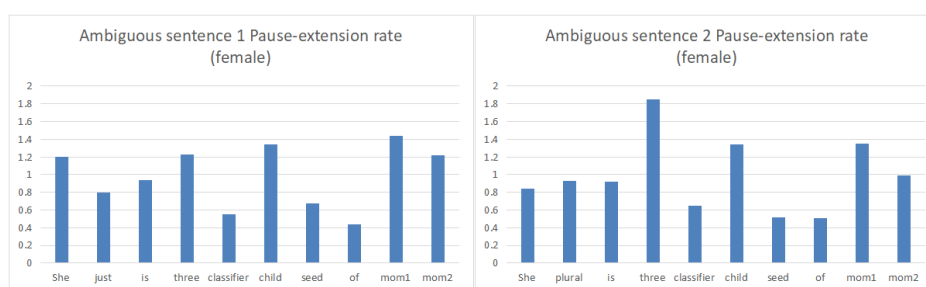
These patterns hold at both the syllable and word/phrase levels: panels with thin boxes reflect syllable-level percentage ranges, whereas thick boxes summarize word/phrase-level ranges (see Section 3.4.2). While the cross-gender profiles are broadly similar, effects are not uniform across tokens or speakers; occasional deviations suggest that other cues and local context can modulate how reliably F0 guides interpretation. Taken together, the fluctuation profiles indicate that focus-related F0 redistribution provides a usable cue to disambiguate the phrase “the mother of three children,” with its informativeness varying by sentence frame and speaker realization.

4.2. Duration performance of the ambiguous structure “the mother of three children” under two semantic interpretations

In Praat, duration data can be obtained by segmenting and measuring each syllable’s length from the waveform. The measurement includes the closure and release phases of stops and affricates. The syllable lengthening ratio is calculated using the lengthening rate formula [12]. If the computed value exceeds 1, it indicates segment lengthening; otherwise, it suggests that no lengthening has occurred. The lengthening rate values for ambiguous sentences 1 and 2, grouped by gender, are shown in Figure 2 and Table 2:



2a. Male Speakers – Ambiguous Sentence 1 (Left)
 2b. Male Speakers – Ambiguous Sentence 2 (Right)



2c. Female Speakers – Ambiguous Sentence 1 (Left)
 2d. Female Speakers – Ambiguous Sentence 2 (Right)
 Figure 2. Lengthening rate patterns by gender (syllables)

From Figure 2, it can be seen that the trends for the male and female groups' sentence structures are generally consistent. Both Ambiguous Sentence 1 and Ambiguous Sentence 2 show pronounced peaks at “three,” “child,” and “mother.” However, there is no significant variation in the dwell time of “child” across the different ambiguous sentences. Therefore, the dwell times for the words “three classifiers” and “mother” in Ambiguous Sentence 1 and Ambiguous Sentence 2 were further calculated for each gender group, with the values presented in Tables 1 and 2 below:

Table 1. Female dwell time (words)

	“Three classifiers”	“mother”
She is the mother of three children. (Ambiguous sentence 1)	0.89	1.33
They are mothers of three children. (Ambiguous sentence 2)	1.25	1.07

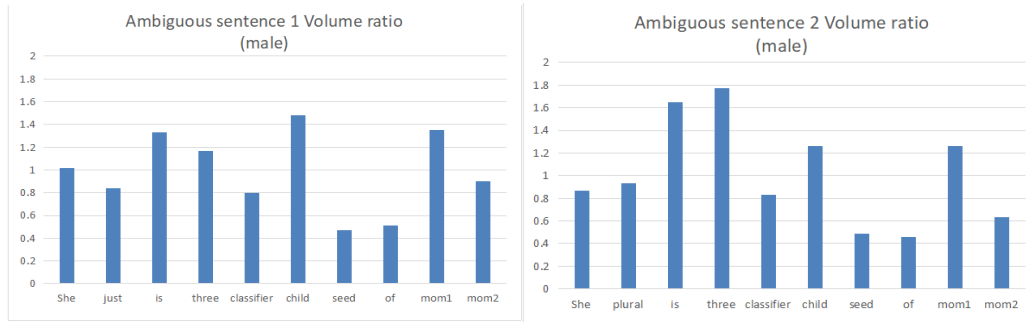
Table 2. Male dwell time (words)

	“Three classifiers”	“mother”
She is the mother of three children. (Ambiguous sentence 1)	0.99	1.17
They are mothers of three children. (Ambiguous sentence 2)	1.28	1.05

From Tables 1 and 2, it can be observed that in Ambiguous Sentence 1, the dwell time for “mother” is greater than for “three classifiers,” and it has the longest dwell time in the entire sentence. In Ambiguous Sentence 2, the dwell time for “three classifiers” exceeds that of “mother,” making it the word with the longest dwell time in that sentence. Words with higher dwell times have longer sound durations and are more prominent. Thus, it can be inferred that in terms of sound duration, “mother” in Ambiguous Sentence 1 has a longer sound duration, while “three classifiers” in Ambiguous Sentence 2 has a longer sound duration. Therefore, in these two ambiguous sentences, “mother” is emphasized in Ambiguous Sentence 1, and “three classifiers” is emphasized in Ambiguous Sentence 2. No significant anomalies were observed in the data. In Ambiguous Sentence 1, “mother” has a higher dwell time in both gender groups, indicating a more significant elongation of its sound segment, and thus it is more prominent. In Ambiguous Sentence 2, “three classifiers” has a higher dwell time, with its sound segment elongated more noticeably, making it more attention-grabbing, consistent with the general correlation between sound duration and prominence.

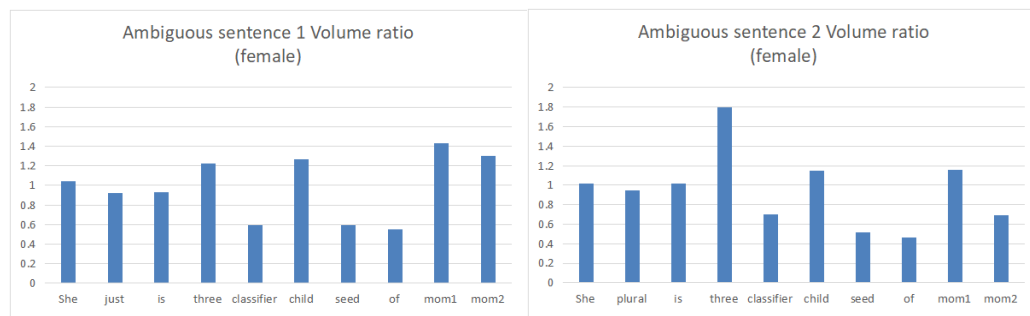
4.3. Acoustic intensity representation of the ambiguous structure “the mother of three children” in two different meanings

In addition to pitch and duration, sound intensity also plays a crucial role in observing fluctuations and stress patterns in sentences, making it an important aspect of prosodic research. To examine the distribution of sound intensity for individual words with different meanings, the formula for calculating volume ratio was used, and normalization was applied. The volume ratio values for the gender groupings in Ambiguous Sentence 1 and Ambiguous Sentence 2 are shown in Figure 3:



3a. Male Ambiguous Sentence 1 (Left)

3b. Male Ambiguous Sentence 2 (Right)



3c. Female Ambiguous Sentence 1 (Left)

3d. Female Ambiguous Sentence 2 (Right)

Figure 3. Volume ratio by gender (word)

From Figure 3, it can be seen that the trends for the male and female groups’ sentence structures are generally consistent. Both Ambiguous Sentence 1 and Ambiguous Sentence 2 show pronounced peaks at “three,” “child,” and “mother,” but the volume ratio for “child” does not change significantly across the different ambiguous sentences, meaning it does not affect our understanding of the two ambiguous sentences. The volume ratios for the words “three classifiers” and “mother” in Ambiguous Sentence 1 and Ambiguous Sentence 2 were further calculated for each gender group, with the values presented in Tables 3 and 4 below:

Table 3. Female volume ratio (words)

	“Three classifiers”	“mother”
She is the mother of three children. (Ambiguous sentence 1)	1.8	2.7
They are mothers of three children. (Ambiguous sentence 2)	2.5	1.85

Table 4. Male volume ratio (words)

	“Three classifiers”	“mother”
She is the mother of three children. (Ambiguous sentence 1)	2.1	2.3
They are mothers of three children. (Ambiguous sentence 2)	2.6	1.9

According to Tables 3 and 4, in these two ambiguous sentences, the words “three classifiers” and “mother” have higher volume ratios and are emphasized. Overall, there is a consistent pattern: these two words generally have higher volume ratios in sentences with different gender and meanings. However, there are special cases. In Ambiguous Sentence 2, the volume ratio for “mother” in the male group is 2.3, lower than the female group’s 2.7, which contradicts the usual assumption that males generally have louder volumes. This could be because, in actual pronunciation, the influence of semantic understanding and expression habits on volume exceeds that of gender factors. In this case, the main role of the volume ratio is to emphasize key semantic information through volume contrast, while the absolute volume size does not play a decisive role in prominence.

4.4. Prosodic focus

Through the above analysis, the experimental results for the male and female groups were averaged, and the data for Ambiguous Sentence 1 and Ambiguous Sentence 2 were summarized in terms of pitch, duration, and intensity. In Ambiguous Sentence 1, “mother” has a higher pitch variation than “mother” in Ambiguous Sentence 2, and in terms of both duration and intensity, it is longer than other prosodic words in the sentence, such as “three classifiers” and “child,” both in dwell time and volume ratio. Therefore, “mother” can be considered the prosodic focus of this ambiguous sentence, which is understood as “a mother has three children.”

In Ambiguous Sentence 2, the numeral–classifier unit “sān gè” (“three”) shows the highest F0 fluctuation (Q_x), greater duration ratio (D_x/D'_x), and a higher volume ratio (E_x) than the other prosodic words in the sentence. On this basis, “sān gè” functions as the prosodic focus for this frame. The location of this focus is consistent with the “three mothers with children” interpretation. Moreover, the focus-based reading is directionally aligned with conclusions drawn from hierarchical, semantic-pointing, and transformational analyses of the same pattern. While individual tokens and speakers occasionally deviate, the aggregate profiles indicate that prosodic measurements (Q_x , D_x/D'_x , E_x) can reliably bias the interpretation of this ambiguous structure toward the reading encoded by the realized focus.

5. Perception and discrimination experiment on the ambiguous structure of “the mother of three children” under different meanings

To further verify the differences between the various parameters reflected in the above acoustic experiments of the ambiguous structure of “the mother of three children,” and to explore whether these differences—such as pitch, pauses, and stress patterns—can be accurately perceived and received by listeners in practical spoken language, a perception and discrimination experiment is designed to study the two different meanings derived from this ambiguous structure.

5.1. Experiment design

5.1.1. Participants

This study surveyed 10 undergraduate students, 5 males and 5 females, all aged 20, from a northern dialect region, with normal hearing.

5.1.2. Experiment corpus

Four audio files were selected based on the most accurate understanding of the two meanings from the acoustic experiment, with the pronunciation closest to the experimental data.

5.1.3. Methodology

The 4 collected audio files were randomly ordered. According to the acoustic experiment results, the perceptual focus of ambiguous sentence 1 should be “mother,” and the perceptual focus of ambiguous sentence 2 should be “three classifiers.” Listeners were asked to report their understanding based on what they heard and to restore it to the different semantic contexts. The results were then statistically analyzed and graphed.

5.2. Discussion of experiment results

The statistical analysis of the listeners’ perception results under the two meanings is summarized in the following table. See Table 5:

Table 5. perception experiment restoration results

Investigate the number of sentence restoration times	“mother”	“Three classifiers”	Correct restoration rate
She is the mother of three children. (Ambiguous sentence 1)	10	0	100%
They are mothers of three children. (Ambiguous sentence 2)	0	10	100%

From the data in Table 5, it can be seen that when the prosodic focus of “the mother of three children” is on “mother,” all 10 listeners correctly restored it, achieving a 100% restoration rate. Similarly, when the prosodic focus was on “three classifiers,” the restoration rate was also 100%. This indicates that listeners can accurately perceive the acoustic emphasis in each semantic context, effectively capturing the intrinsic differences in stress characteristics exhibited by the ambiguous structure “the mother of three children” under different meanings. Thus, listeners can use speech information to decode the meaning and restore the correct interpretation.

6. Conclusion

Through the analysis of the prosodic focus performance of the ambiguous structure “the mother of three children” and the investigation of the prosodic distribution patterns in terms of pitch, duration, and intensity, the following conclusions are drawn:

Acoustic Experiment Findings:

Based on the acoustic experiment, the prosodic focus of the ambiguous structure “the mother of three children” was analyzed, with a focus on the distribution patterns of pitch, duration, and intensity. The results show that under different meanings, the pitch peak of the prosodic focus significantly rises, the duration clearly extends, and the intensity also exhibits an increasing trend. Thus, the prosodic focus plays a crucial role in the semantic interpretation of this ambiguous structure, and its

acoustic characteristics serve as effective markers for distinguishing the ambiguity. The study further validates the role of prosodic features in syntactic ambiguity differentiation and provides experimental phonetic evidence for the study of Chinese prosodic focus and ambiguity resolution.

Perception experiment results

Listeners systematically exploited acoustic differences associated with the two readings of “the mother of three children.” In aggregate, higher F0 fluctuation (Qx) and volume ratio (Ex) on the numeral–classifier unit, together with an elevated duration ratio (Dx/D’x), biased responses toward the “three mothers with children” interpretation; the reverse prominence pattern favored the “one mother of three children” reading. Although effect sizes varied across tokens and participants, the direction of bias matched the realized focus in the production data.

Interdependence of prosodic elements

The three cues—pitch (Qx), duration (Dx/D’x), and intensity (Ex)—contributed jointly rather than identically. Duration and intensity tended to pattern consistently with focus placement across speakers, while pitch was more discriminative but also more variable across sentence frames; identifying the focus therefore benefited from comparative cue evaluation instead of relying on any single parameter.

Implications and next steps

These results indicate that focus-related prominence provides usable perceptual evidence for disambiguating this structure. To refine boundary conditions, future work should (i) test cue-weighting with controlled cue conflict designs, (ii) examine contextual interactions (e.g., discourse givenness), and (iii) extend the paradigm to additional ambiguous patterns with comparable production–perception coupling.

References

- [1] Feng, X. M. (2007). Analysis methods of ambiguous sentence patterns in modern Chinese. *Hebei Self - study Examinations*, (4), 6.
- [2] Ma, C. Y. (2013). The gradual emergence of semantic factors in structural analysis - from the sentence - component analysis method to the sentence - pattern transformation analysis method. *Journal of Xingyi Normal University for Nationalities*, (6), 50-53.
- [3] Zhou, G. (1998). A tentative discussion on semantic orientation analysis. *Linguistic Research*, (3), 27-34.
- [4] Zhu, X. N. (2024). The four - stage development of phonetics. *Journal of Changshu Institute of Technology*, 38(1), 24-46. DOI: 10.16101/j.cnki.cn32 - 1749/z.2024.01.010
- [5] Luo, L. (2022). An acquisition study on the prosodic - related interfaces of Chinese ambiguous structures by Uyghur students. [Doctoral dissertation, Sichuan International Studies University]. DOI: 10.27348/d.cnki.gscwc.2022.000022
- [6] Xue, L. Y. (2015). The use of prosodic cues in second - language spoken language processing. [Doctoral dissertation, Zhejiang University].
- [7] Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *The Journal of the Acoustical Society of America*, 53(1_Supplement), 380.
- [8] Halliday, M. A. K. (1967). Notes on transitivity and theme in English: Part 2. *Journal of Linguistics*, 3(2), 199-244.
- [9] Huang, C. Y. (2012). An experimental phonetic analysis of the “V_dual + N_dual” ambiguous structure. *Language Teaching and Linguistic Studies*, (3), 98-104.
- [10] Wen, B. Y., & Dong, W. Y. (2019). Analysis of the prosodic focus of the “lian NP dou VP” sentence pattern. *Chinese Language Learning*, (5), 26-36.
- [11] Hao, C. X., & Lü, M. (2023). Research on the prosodic focus of disambiguating the “V - de - A” structure from the perspective of experimental phonetics. In *School of Chinese as a Second Language, Peking University (Eds.), Proceedings of the 2023 Doctoral Forum on Teaching Chinese as a Foreign Language and the 16th Academic Forum for Postgraduate Students in Teaching Chinese as a Foreign Language* (pp. 12). [Publisher not specified]. DOI: 10.26914/c.cnkihy.2023.022038
- [12] Shi, F., Wang, P., & Liang, L. (2009). The pitch range of declarative intonation in Mandarin Chinese. *Nankai Linguistics Journal*, (2), 4-13, 178.

- [13] Liang, L., & Shi, F. (2010). Analysis of the volume ratio of two - character groups in Mandarin. *Nankai Linguistics Journal*, (2), 35-41, 186.