



Influences of Morphosyntax and Semantics on the Intonation of Mandarin Chinese *Wh*-indefinites

Hongchen Wu¹, Jiwon Yun²

¹Georgia Institute of Technology, USA

²Stony Brook University, USA

hwu480@gatech.edu, jiwon.yun@stonybrook.edu

Abstract

This study utilized a speech production test to examine how different components of the language faculty interact, focusing on the impact of morphosyntax and semantics on the intonation of Mandarin Chinese *wh*-indefinites. The analysis of the hour-long speech production data from 33 speakers indicates that both morphosyntax and semantics play a crucial role in shaping the intonation of *wh*-indefinites. We found that *wh*-indefinites in complex clause structures effectively utilize all prosodic features to convey interrogative meaning. Moreover, when a morphosyntactic marker suggests a preference for a specific interpretation of *wh*-indefinites, there is a reduced prosodic contrast (although still significantly different) between *wh*-interrogatives and *wh*-indefinites, underlying an interference effect of morphosyntactic cues in speech production.

Index Terms: *wh*-prosody; indefinite readings; interrogative readings; speech production

1. Introduction

One of the fundamental questions in linguistics is how different components of the language faculty interact with each other. Yet, systematic investigations into this question have been limited, partially because a controlled paradigm to study linguistic interfaces is not readily available. An emerging effort to contribute to this fundamental inquiry is to investigate the linguistic interfaces related to *wh*-indefinites. Words that can be used to form content questions are referred to as *wh*-words since such words in English (e.g., *who* and *what*) often start with *wh*-. *Wh*-words in many languages can alternate between an indefinite and an interrogative reading, thus are referred to as *wh*-indefinites [1, 2]. For example, the Mandarin Chinese *wh*-word *shenme* can be interpreted as either ‘something’ or ‘what’, rendering the corresponding sentence type as a statement or a question.

Numerous studies have investigated the distinction between the two readings of *wh*-indefinites, primarily through their distinctive intonation patterns (e.g., Japanese: [3], Korean: [4], Hong Kong Cantonese: [5], Mandarin Chinese Putonghua: [6], Taiwan Mandarin: [7, 8]). Across languages, a common finding is that *wh*-interrogatives are prosodically more prominent than *wh*-indefinites. However, previous studies investigating the interaction between speech production and syntax-semantics of *wh*-indefinites have largely focused on simple sentences, thus it remains unclear whether similar patterns persist within more complex clause structures. Moreover, studies examining the prosodic prominence of *wh*-interrogatives compared to *wh*-indefinites in Mandarin Chinese have yielded conflicting results regarding the specific prosodic features associated with

wh-prominence in interrogatives. While some emphasize pitch prominence exclusively [9, 10, 11, 12], others also highlight duration or intensity prominence [6, 13, 14]. These discrepancies may stem from the imbalanced positioning of *wh*-words within target sentences and the utilization of non-identical strings for *wh*-interrogatives and *wh*-indefinites. Additionally, many of these studies relied on relatively small sample sizes (4 to 10 speakers), often exhibiting unbalanced gender representation and lacking pitch normalization during analysis. These limitations may hinder the ability to draw general conclusions regarding the intonation of *wh*-indefinites.

The present study contributes to the linguistic inquiry on how the language faculty interacts by examining the influences of morphosyntax and semantics on the intonation of Mandarin Chinese *wh*-indefinites embedded in conditional clauses. These clauses provide complex yet identical environments for interrogative and indefinite interpretations. In Mandarin, conditional clauses are generally marked by *ruguo* and *dehua*, either individually or in combination, while it is possible to construct a conditional clause without any morphological marking (Table 1). *Ruguo* emphasizes hypotheticality and *dehua* emphasizes conditioning, and the two morphological markers collectively account for about 94% of conditional occurrences in Mandarin Chinese [15]. Although conditional clauses can be a licensing environment for *wh*-indefinites (e.g., [16]), it should be noted that each of the given sentences in Table 1 could potentially have two readings.

Table 1: Examples including *wh*-indefinites in Mandarin Chinese conditional clauses

| | <i>Ruguo</i> if | <i>Wh</i> what | <i>Verb</i> enter | <i>Object</i> reagent | <i>Dehua</i> in-case | <i>Then</i> dangerous |
|----|--------------------|-------------------|----------------------|--------------------------|-------------------------|--------------------------|
| a. | 如果 | 什么 | 掉进 | 碗里 | | 会很危险 |
| b. | | 什么 | 掉进 | 碗里 | 的话 | 会很危险 |
| c. | 如果 | 什么 | 掉进 | 碗里 | 的话 | 会很危险 |
| d. | | 什么 | 掉进 | 碗里 | | 会很危险 |

(i) *wh*-indefinite: ‘If something falls into that chemical reagent, it will be dangerous.’

(ii) *wh*-interrogative: ‘What is the thing such that it would be dangerous if it falls into that chemical reagent?’

With this controlled design, which provides identical strings for *wh*-indefinites and *wh*-interrogatives, we conducted a production study to identify the influences of morphosyntax (i.e., morphological markers for conditional clauses) and semantics (i.e., the readings of *wh*-indefinites) on the intonation of *wh*-indefinites in speech production. Based on existing knowledge of the intonation of *wh*-indefinites,

two hypotheses and their corresponding predictions emerge. Firstly, prosodic prominence of *wh*-words corresponds to *wh*-interrogative readings. We predict that the *wh*-region will consistently exhibit prosodic prominence, whether through pitch prominence, longer duration, or higher intensity, for *wh*-interrogatives. This contrastive pattern is expected to hold irrespective of morphosyntactic clues, such as the specific morphological marker used for conditional clauses. Secondly, morphosyntactic markers for conditional clauses favor an indefinite reading. We thus anticipate differences in the phonetic realization of the intonation patterns of *wh*-interrogatives depending on the specific morphological marker used for conditional clauses.

2. Methodology and Dataset

We controlled for the morphological markers for conditional clauses (the presence and absence of *ruguo* and *dehua*) and the meaning of the target sentences (indefinite versus interrogatives) in our stimuli. Therefore, Morphology and Meaning serve as the two independent variables in our study, predicting intonation of *wh*-indeterminates, which is measured in pitch, intensity, and duration, as the dependent variable. Additionally, we balanced the animacy of the *wh*-indeterminates (*shenme* ‘what’ versus *shenme ren* ‘what people’) and the position of *wh*-indeterminates (subject versus object position) in the stimuli to introduce lexical and positional diversity. 16 sentences like those in Table 1 were created, and each sentence is paired with two possible readings: one is a *wh*-indefinite reading, and the other is a *wh*-interrogative reading. The full list of the target sentences can be found in the Supplementary Materials included on <https://github.com/hongchen0926/wh-indeterminates>.

Thirty-four Mandarin adult speakers (age: 19-32) were recruited through emails and social media to participate voluntarily in this speech production experiment. During the recording procedure, they were instructed to articulate either the designated target sentence based on the given meaning, or indicate that they believed the target sentence did not accurately convey the intended meaning if they perceived it as such. Participants were not restricted in the timing of their recordings or the number of attempts to capture the target sentences. They had the flexibility to update their recording samples at any time by re-recording a sentence before submission, with only their final audio recording version being included in the dataset. Data from one participant was excluded due to incomplete recordings.

Out of the 33 participants whose recordings were included in the final dataset, 18 participants (9 females and 9 males) recorded their speeches in a soundproof booth using professional recording equipment. The other 15 participants (8 females and 7 males) recorded their speeches using either their smartphones or laptops in a quiet room and submitted their speech recordings via email. We converted the non-lab recordings into the .wav format using Praat (version 6.1.53, [17]) if they were in submitted a different format. The data collection approach employed in the present study not only allowed us to create relatively larger data samples from more speakers than in previous studies but also allowed us to establish a well-balanced dataset comprising audio files from both laboratory recordings and non-laboratory recordings, with an equitable distribution of files from both female and male speakers. This balanced representation helps to minimize potential biases stemming from variations in audio input quality and gender influences on speech productions.

All 33 participants accepted the dual interpretations of the majority of sentences in the stimuli, indicating that *wh*-indeterminates in the given sentences are indeed ambiguous for these speakers. Each participant contributed 32 recorded audio files (16 target sentences \times 2 possible readings) to the dataset, totaling 1,056 audio files. These audio files have a total duration of 1 hour and 2 seconds. After removing 100 audio files indicating rejection of a specific reading and 1 audio file where the recorded sentence did not match the target sentence, we annotated the rest 955 audio files (totaling 53 minutes and 45 seconds) using the Montreal Forced Aligner [18] and the workflow designed in [19] with subsequent human adjustment, yielding 5,331 data entries at the end.

3. Results

We measured the maximum pitch heights, pitch ranges, duration, and maximum intensity of the *wh*-region in each audio file. We subsequently normalized both the maximum pitch values and the pitch excursion values using the z-score, which helps us ensure that any observed differences in pitch are not simply due to individual differences in pitch range or variability among speakers.

To ensure data quality, any data point with a z-score of maximum pitch or a z-score of pitch excursion greater than 3 or less than -3 was considered an outlier. We additionally eliminated all instances of “N/A” values. Following this filtering process, 93.91% of the data points were retained. We used the *lme4* package (version 1.1-35.1, [20]) in R (version 4.3.2, [21]) to perform a mixed linear regression model to compute inferential statistics. We fitted the linear mixed model to predict prosodic values of *wh*-words with Meaning (the meaning of the sentence: indefinite or interrogative) and Morphology (the morphological marker for conditional clause: *ruguo*, *dehua*, *ruguo...dehua*, or no morphological marker). The model included Participant as a random intercept effect. The final dataset, along with the R code used for the statistical analysis and data visualization, can be found in the Supplementary Materials included on <https://github.com/hongchen0926/wh-indeterminates>.

We first used the *anova()* function to compare an interdependent model that predicts prosodic values using the interaction of morphological markers and meaning of *wh*-word with a non-interdependent model. Results returned by *anova()* tests suggest that the interdependent model does not significantly improve the fit compared to the non-interdependent model (Table 2). Therefore, we fitted a non-interdependent model to predict prosodic values of *wh*-words with Meaning and Morphology as non-interdependent fixed factors.

Table 2: Comparisons between an interdependent model and a non-interdependent model

| Dimensions | anova() results |
|----------------|-----------------------------|
| Max_pitch_Z | $\chi^2(3) = 0.64, p = .89$ |
| Pitch_Ex_Z | $\chi^2(3) = 0.11, p = .99$ |
| Duration_in_ms | $\chi^2(3) = 0.70, p = .87$ |
| Max_dB | $\chi^2(3) = 0.20, p = .98$ |

Regarding the influence of Meaning on the intonation of *wh*-indeterminates, we found that the effect of Meaning [interro] is statistically significant, aligning with prior research on speech production of *wh*-indeterminates [6, 9, 10, 13]. However, our results reveal notable distinctions compared to previ-

ous studies that reported only a few factors active in prosodic disambiguation, with *wh*-interrogatives displaying all-rounded prominence: *wh*-interrogatives have significantly higher maximum pitch, pitch excursion, duration, and maximum intensity at the *wh*-region, compared to *wh*-indefinites (Table 3). These results also confirm our prediction that *wh*-region consistently exhibits prosodic prominence for *wh*-interrogatives, even in complex clause structures, irrespective of the morphosyntactic interference.

Table 3: Influences of Meaning on prosodic values of *wh*-words

| Dimensions | Effect of Meaning [interro] |
|----------------|--|
| Max_pitch_Z | $\beta = 0.50$, SE = 0.04, $p < .001$ |
| Pitch_Ex_Z | $\beta = 0.52$, SE = 0.04, $p < .001$ |
| Duration_in_ms | $\beta = 119.19$, SE = 8.13, $p < .001$ |
| Max_db | $\beta = 2.20$, SE = 0.19, $p < .001$ |

Regarding the effect of Morphology on the intonation of *wh*-indefinites, we observed that the four types of morphological conditional markers emerge into two categories: one with *ruguo* and one without *ruguo*. Table 4 shows that using Morphology [none] as the baseline, the effect of Morphology [dehua] is found to be statistically non-significant. However, both Morphology [*ruguo*] and Morphology [*ruguo...dehua*] exhibit statistically significant and negative effects, particularly concerning the maximum pitch, pitch excursion, and maximum intensity of *wh*-indefinites. In other words, employing *ruguo* as the conditional marker leads to a significant reduction in the contrast between *wh*-interrogatives and *wh*-indefinite, as illustrated in Figure 1.

Table 4: Influences of Morphology on prosodic values of *wh*-words

| | Effect of Morphology [<i>dehua</i>] |
|----------------|---|
| Max_pitch_Z | $\beta = 0.04$, SE = 0.05, $p = .46$ |
| Pitch_Ex_Z | $\beta = 0.01$, SE = 0.05, $p = .85$ |
| Duration_in_ms | $\beta = 20.21$, SE = 11.45, $p = .08$ |
| Max_db | $\beta = 0.36$, SE = 0.26, $p = .16$ |
| | Effect of Morphology [<i>ruguo</i>] |
| Max_pitch_Z | $\beta = -0.20$, SE = 0.05, $p < .001$ |
| Pitch_Ex_Z | $\beta = -0.11$, SE = 0.05, $p = .03$ |
| Duration_in_ms | $\beta = -6.17$, SE = 11.39, $p = .59$ |
| Max_db | $\beta = -0.74$, SE = 0.26, $p = .005$ |
| | Effect of Morphology [<i>ruguo...dehua</i>] |
| Max_pitch_Z | $\beta = -0.21$, SE = 0.05, $p < .001$ |
| Pitch_Ex_Z | $\beta = -0.10$, SE = 0.05, $p = .05$ |
| Duration_in_ms | $\beta = 14.90$, SE = 11.54, $p = .20$ |
| Max_db | $\beta = -0.74$, SE = 0.26, $p = .005$ |

4. Discussions

The results show that *wh*-interrogatives display significantly higher values across all different prosodic measurements compared to *wh*-indefinites at the *wh*-region. This contrasts with the varying degrees of *wh*-prominence for *wh*-interrogatives in earlier studies that focused on the speech production of simple sentences containing Mandarin *wh*-indefinites. Some studies

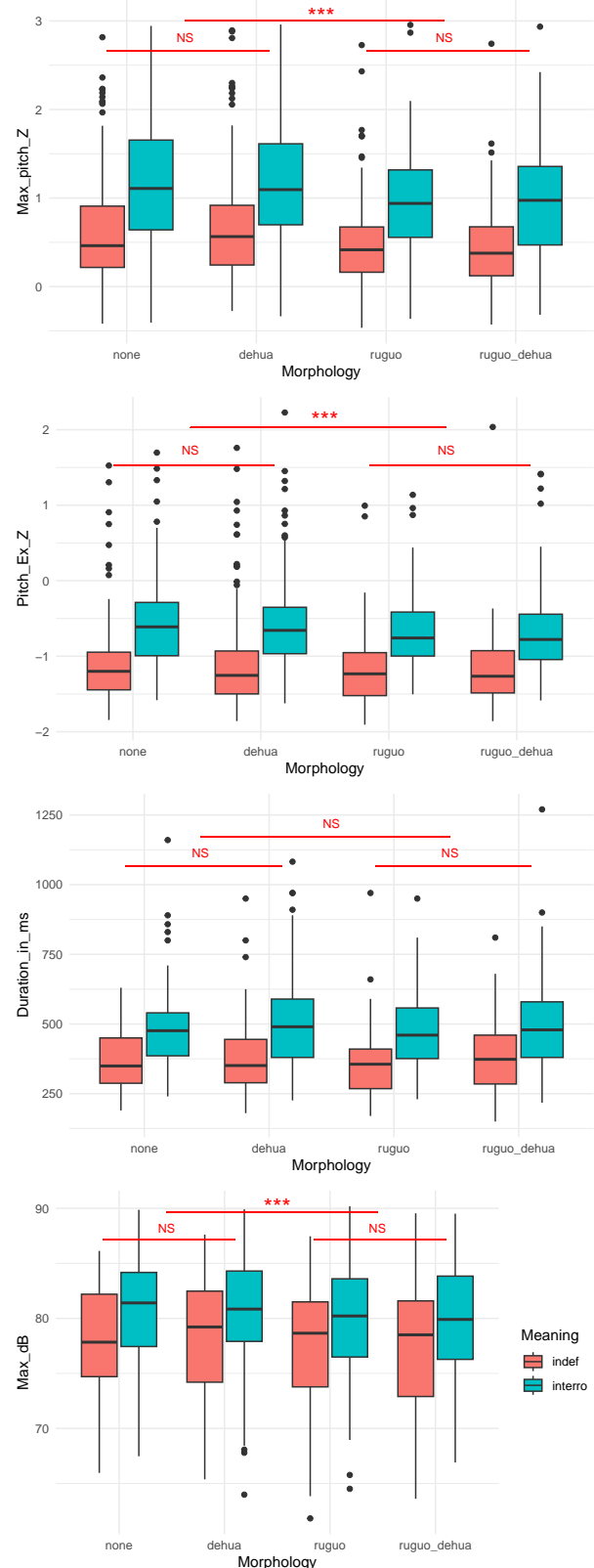


Figure 1: Measurement of acoustic properties at the *wh*-region: normalized maximum pitch, normalized pitch excursion, duration, maximum intensity

such as [9, 10, 12] only observed pitch prominence but not other prosodic prominence in *wh*-interrogatives. Conversely, other studies such as [7] and [14] only observed other prosodic prominence (longer duration, greater intensity) for *wh*-interrogatives but not pitch prominence. On the other hand, research by [6, 13] revealed pitch prominence, longer duration, and higher intensity in *wh*-interrogatives within the *wh*-region, but it is difficult to draw a general conclusion as these studies only surveyed *wh*-words in the object position. The present study is one of its first kind, that found *wh*-prominence across multiple prosodic measurements including pitch peak, pitch excursion, duration, and intensity, when the *wh*-indeterminate is an argument of a simple sentence. Our results present an all-inclusive pattern of *wh*-prominence, suggesting that *wh*-indeterminates in a complex clausal environment (such as conditional clauses) tend to employ all possible prosodic features to signal an interrogative reading effectively.

Regarding the influence of morphological markers, we observed a reduced contrast pattern for conditions involving *ruguo*: *ruguo* and *ruguo...dehua* make the prosodic contrasts between interrogative and indefinite readings significantly smaller than *dehua* and *none* conditions make. We found that such a pattern is more consistent when the *wh*-indeterminate is in the subject position, thus one might wonder if the reduced contrast pattern stems from subject *wh*-words being in sentence-initial position for *dehua* and *none* conditions whereas being in non-sentence-initial position for *ruguo* and *ruguo...dehua* conditions. However, a recent study [14] compared the speech production of subject *wh*-indeterminates in both sentence-initial and non-sentence-initial positions, yet found no significant differences in the intonation at the *wh*-region between the two conditions.

We attribute the reduced contrast pattern to the interference of morphological markers with speech production. As mentioned earlier, while all the morphological markers are compatible with both statements and questions, conditional clauses are one of the canonical environments that can accommodate *wh*-indefinites [16]. When encountering the morphological marker *ruguo* before a *wh*-indeterminate, the readiness of a licensing environment for *wh*-indefinites could lead speakers unconsciously produce *wh*-indefinite-like speech pattern for *wh*-interrogative readings. In other words, with the presence of *ruguo*, the prosodic patterns of *wh*-interrogatives still exhibit significant differences from those of *wh*-indefinites, but they become assimilated with the prosodic patterns of *wh*-indefinites. If speakers were consciously considering a speech strategy to make *wh*-interrogatives more prominent when *ruguo* is present, we would then expect the opposite, increased contrast pattern: *ruguo* conditions would be expected to enhance the interrogative versus indefinite prosodic contrasts significantly more than non-*ruguo* conditions. Therefore, the observation of a reduced contrast pattern suggests that morphosyntax creates an interference effect rather than providing a strategic cue for speech production in this case.

5. Conclusion and future work

The experimental results reveal that both morphosyntax and semantics play a role in shaping the intonation of *wh*-indeterminates. Firstly, *wh*-indeterminates within complex clause structures favor employing all available prosodic features to effectively convey the interrogative interpretation. Secondly, when a morphosyntactic marker indicates the readiness of a licensing environment for *wh*-indefinites, a reduced

prosodic contrast is observed between *wh*-interrogative and *wh*-indefinites, indicating interference from morphosyntax in speech production.

The present study primarily focuses on speech production, and we aim to extend our research to include speech perception to investigate the interfaces between morphosyntax, semantics, and prosody of *wh*-indeterminates from a perceptual perspective. Some of the following-up research questions include investigating which prosodic cues (such as pitch, intensity, duration) can be more predictive of speech perception regarding a particular reading of *wh*-indeterminates. In addition, we aim to explore whether morphosyntax interferes with speech perception in a similar way as it interferes with speech production.

6. Acknowledgements

The experiment protocol was approved by the Institutional Review Board at Georgia Institute of Technology (Protocol Number: H21340) and the Institutional Review Board at Stony Brook University (Protocol Number: IRB 2021-00358) as minimal risk research qualified for exemption status. We sincerely thank all the participants who took the time to participate in the production experiments. Our sincere thanks also go to undergraduate student assistants (Xiang Li, Huiyi Huang, and Chuan-dong Liu) for their help with audio annotation. This research was funded by Faculty Professional Development funds from the School of Modern Languages and Ivan Allen College at the Georgia Institute of Technology, awarded to the first author.

7. References

- [1] S.-Y. Kuroda, "Generative grammatical studies in the Japanese language," Ph.D. dissertation, Massachusetts Institute of Technology, 1965.
- [2] J. Yun, "Meaning and prosody of *wh*-indeterminates in Korean," *Linguistic Inquiry*, vol. 50, no. 3, p. 630–647, 2019.
- [3] S. Ishihara, "Invisible but Audible *Wh*-Scope Marking: *Wh*-constructions and Deaccenting in Japanese," in *Proceedings of the 21st West Coast Conference on Formal Linguistics*, L. Mikkelsen and C. Potts, Eds., 2002, pp. 180–193.
- [4] J. Yun and H.-S. Lee, "Prosodic disambiguation of questions in Korean: Theory and processing," *Korean Linguistics*, vol. 18, no. 1, pp. 18–47, 2022.
- [5] Y.-Y. Hsu and A. Xu, "*Wh*-indeterminates and Prosody in Hong Kong Cantonese," in *Proc. Speech Prosody 2020*, 2020, pp. 376–380.
- [6] —, "Interaction of prosody and syntax-semantics in Mandarin *wh*-indeterminates," *J. Acoust. Soc. Am.*, vol. 148, no. EL119–EL124, 2020. [Online]. Available: <https://doi.org/10.1121/10.0001676>
- [7] S. Shyu and T.-L. Tung, "Mandarin *wh*-phrases and prosody," *Studies in Prosodic Grammar*, vol. 3, pp. 32–76, 2018.
- [8] Y.-Y. Hsu and A. Xu, "Sentence Prosody and *Wh*-Indeterminates in Taiwan Mandarin," in *Proc. Interspeech 2019*, 2019, pp. 3950–3954.
- [9] F. Hu, "A Prosodic Analysis of *Wh*-words in Standard Chinese," in *Proceedings of Speech Prosody*, 2002, pp. 403–406. [Online]. Available: https://www.isca-speech.org/archive/speechprosody_2002/02_hu02_speechprosody.html
- [10] H. Dong, "Issues in the Semantics of Mandarin Questions," Ph.D. dissertation, Cornell University, 2009. [Online]. Available: <https://hdl.handle.net/1813/13930>
- [11] X. Liu, A. Li, and Y. Jia, "How does prosody distinguish *Wh*-statement from *Wh*-question? A case study of Standard Chinese," 05 2016, pp. 1076–1080.

- [12] P. Wang and X. Wang, “Comparing the prosodic and perceptual characteristics of the Chinese interrogative pronoun *Shui when* indicating inquiry and free reference,” *Chinese Teaching in the World*, vol. 3, pp. 379–391, 2020.
- [13] Y. Yang, “The Two Sides of *Wh*-Indeterminates in Mandarin: A Prosodic and Processing Account,” Ph.D. dissertation, Leiden University, Leiden, Netherlands, 2018. [Online]. Available: <https://hdl.handle.net/1887/62454>
- [14] H. Wu and J. Yun, “Holistic Prosodic Examination of Mandarin *Wh*-indeterminates,” 2023, work in progress, manuscript.
- [15] H.-C. Yeh, “Temporal and Conditional Clauses in Chinese Spoken Discourse: A Function-based Study,” in *Proceedings of the 14th Pacific Asia Conference on Language, Information and Computation*, A. Ikeya and M. Kawamori, Eds. Tokyo, Japan: PACLIC 14 Organizing Committee, Waseda University, 2000, pp. 365–376. [Online]. Available: <https://aclanthology.org/Y00-1036.pdf>
- [16] Y.-H. A. Li, “Indefinite *Wh* in Mandarin Chinese,” *Journal of East Asian Linguistics*, vol. 1, no. 2, pp. 125–155, 1992. [Online]. Available: <http://www.jstor.org/stable/20100614>
- [17] P. Boersma and D. Weenink, “Praat: Doing phonetics by computer [Computer program],” Version 6.1.53, 2021. [Online]. Available: <http://www.praat.org/>
- [18] M. McAuliffe, M. Socolof, S. Mihuc, M. Wagner, and M. Sonderegger, “Montreal Forced Aligner: trainable text-speech alignment using Kaldi,” in *Proceedings of the 18th Conference of the International Speech Communication Association*, 2017, pp. 498–502. [Online]. Available: <https://doi.org/10.21437/Interspeech.2017-1386>
- [19] H. Wu, J. Yun, X. Li, H. Huang, and C. Liu, “Using a forced aligner for prosody research,” *Humanities and Social Sciences Communications*, vol. 10, p. 429, 2023. [Online]. Available: <https://doi.org/10.1057/s41599-023-01931-4>
- [20] D. Bates, M. Mächler, B. Bolker, and S. Walker, *lme4: Linear Mixed-Effects Models using Eigen and S4*, 2023, r package version 1.1-35.1. [Online]. Available: <https://CRAN.R-project.org/package=lme4>
- [21] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, 2023, r package version 4.3.2. [Online]. Available: <https://www.R-project.org/>