

Suprasegmental cues and intra-dialectal hierarchies: evidence from northeast mandarin

Yuchen Chu

School of Philosophy, Psychology and Language Sciences (PPLS), University of Edinburgh, Edinburgh, UK

S28061552@ed.ac.uk

Abstract. This research investigates intra-dialectal hierarchies within Northeastern Mandarin, focusing on the Shenyang and Jinzhou dialects, two closely related varieties in Liaoning province, China. The segmental features of these dialects are largely comparable; however, their suprasegmental characteristics, especially the intonation patterns in interrogatives, demonstrate considerable divergence. This enables us to examine how listeners utilize prosodic cues for both recognition and social assessment. The study, which involved recordings of speech, perception tests, and attitude surveys with ninety individuals from both local and non-local backgrounds, reveals a paradox: individuals struggle to accurately identify dialect origins through suprasegmental features, yet consistently evaluate Shenyang speech more favorably, indicating its status as the regional standard. This "misrecognition paradox" asserts that suprasegmental cues can sustain symbolic hierarchies even in the absence of accurate recognition, thus clarifying the implicit mechanisms that contribute to linguistic inequality. The results enhance sociophonetics and sociolinguistics by demonstrating how prosodic features facilitate intra-dialectal stratification and perpetuate social hierarchies beyond overt language classification.

Keywords: dialect recognition, language attitudes, Chinese dialects, sociophonetics, suprasegmental features

1. Introduction

Language, the main vehicle of communication, embodies power and identity in multi-lingual or multi-dialectal context. Linguistic hierarchy that has long been under the spotlight of social-linguistics elaborates on the bi-polarity between highly recognized well-received standard languages and dismissed marginalized non-standard languages. This phenomenon is not exclusive to China; for example, language varieties associated with various regions or social classes also show clear hierarchical stratification in the United States (Lippi-Green, 2012) and the United Kingdom (Trudgill, 1974).

In China's linguistic hierarchy, mandarin, predominant in education, media and social mobility, is universally acknowledged as an orderly and decent linguistic form, to which dialects are inferior, often dismissed as low-brow or non-standard. However, the power inequality may go beyond standard language and dialects and exist between variants of a particular dialect. An intra-dialectal region is not linguistically homogenous, as variants in the central cities better positioned to political, economic and cultural resources are superior to those in subsidiary cities, excluded from the mainstream communication system.

Targeted at Shenyang and Jinzhou dialects, two variants of northeast Mandarin dialect in the mid-Liaoning province, the essay explores intra-dialectal hierarchy and its reflection on linguistic recognition and social comment. In the linguistic system, the two variants, different in dialectal subregions, are similar in segmental structure (Wen, 2012; Zhang, 2014). Nonetheless, a sharp contrast between them falls on social status and media circulation, as the Shenyang dialect originates from Shenyang, the capital city of Liaoning province, widely circulates in multi-media context, and evolves into a representative accent in northeast China, whereas Jinzhou dialect is rarely seen in mainstream communication, much less social visibility. Another sensible contrast lies in such suprasegmental features as intonation contour. Of particular note is the end of interrogative sentences where the intonation of Jinzhou dialect rises, but that of Shenyang dialect remains flat or fall. As a result, it offers an ideal setting for investigating the function of suprasegmental variables in social evaluation and linguistic attitudes, as well as testing whether they can function as reliable perceptual cues in situations where segmental differences are minimized. Therefore, the essay rolls out a comparative analysis on the viability of adopting such suprasegmental difference as an effective sensor, despite segmental similarity, and on the role of suprasegmental features in language attitude and social comment, with a view to unveiling the hierarchy of intra-dialectal variants and its foundation for social perception. By examining how listeners use suprasegmental features for social evaluation in contexts where segmental differences are minimal, this study contributes to our understanding of

hierarchical mechanisms within Northeastern Mandarin while also offering a novel case for cross-linguistic sociophonetician research.

2. Literature review

2.1. hierarchy of intra-dialectal variants

Recent years have seen sociolinguistics shift its focus to the power hierarchy of intra-dialectal variants and its ideological mechanism, apart from the beaten approach of dualism between standard language and dialects. Once "interior standardization" was proposed, Auer (2005) considered it possible for a region in the absence of an official language standard to develop a variant of its central city into an internally "representative" language based on political, economic, and cultural status. This is consistent with Milroy's (2002) claim that language standardization is a social construct rather than a natural process; Silverstein's (2018) concept of the "monoglot standard ideology" further illustrates how institutionalized norms serve to legitimize and normalize linguistic hierarchies. These studies indicate that, justified by media reproduction, distribution of educational resources, among other methods, the hierarchy that features the superiority of central city dialect takes shape and eventually bears on language users' perception and evaluation of the variants.

A case in point is the region of northeast mandarin. While Jinzhou and other non-central variants are still largely marginalized, the Shenyang accent, which is the central city, has progressively been shaped in the media and public opinion as the representative form of Northeastern Mandarin. As is further revealed in the linguistic capital theory of Bourdieu (1991), the standardization of the central-city variant builds up and reproduces linguistic capital, which triggers a vicious circle that further hampers the application of other marginalized variants. Meanwhile, he also coined the term "misrecognition" to describe people's propensity to unwittingly accept and replicate symbolic power, which normalizes and validates linguistic hierarchies.

In Chinese contexts, Vast arrays of empirical research have testified to the awareness among language users of linguistic hierarchy. A diachronic survey conducted by Yan et al. (2024) in Nanchang province concluded that, driven by pro-Mandarin policy, education, and cross-generation legacy, Mandarin prevailed over local dialects and served as the reigning language. Another research on tertiary institutes in Sichuan province undertaken by Xin (2006) illustrated the disequilibrium between mandarin, rated as a prerequisite for social mobility, and local dialects, tangled with such derogative images as discrimination against accents and their origins. Powerful as Shanghaiese and Cantonese dialects are, mandarin's circulation in the two regions reshapes the local people's linguistic attitude. In the light of the empirical research by Zhou (2001), the surveyed Shanghaiese and Cantonese speakers with deep attachment to their mother tongue dialects opted into mandarin on educational, vocational and other occasions where languages were used functionally, for mandarin was entitled to greater linguistic capital than the local dialects. It testified that in a linguistic hierarchy, empathy with the sidelined cannot hold back the institutional leadership of those at the center.

However, vast majority of aforesaid research centers on the formation of hierarchy between mandarin and local dialects, as opposed to the differentiated attitude towards variants of a particular dialect. Under the framework of the theories above, the essay sheds some light on the formation and maintenance of intra-dialectal hierarchy based on symbolic comments in a non-standard linguistic context and evidenced by empirical research on recipients' language attitude and evaluation between Shenyang and Jinzhou dialects in and outside northeast China. In addition to filling the gaps in domestic research, this also speaks to global sociolinguistic debates about internal hierarchies, ideology, and standardization.

2.2. The impact of suprasegmental features on dialectal recognition and attitude

Conventional research on dialectal recognition have laid greater emphasis on such remarkable segmental features as initials, finals, syllable structure and tone system between mandarin and local dialects, including the Wu dialect and Southern Fujian dialect (Zheng, 2007; Li et. al, 2006), which unveils segmental differences are of great significance to speech recognition across dialectal regions. Nonetheless, the recognition based on segmental differences falls short of precision, so research are in need of higher-precision sensors.

Against the backdrop, academia shifts its focus on suprasegmental features, including such rhythmic variables as intonation, pitch, rhythm, stress and speed, in search for potential sensors in macro-linguistic structure. According to preliminary research by Clopper and Pisoni (2004), suprasegmental features serve as key perceptive sensor in cases where segmental information is limited or insufficient. It is further verified in a series of research ranging from speech perception, phonics modelling, to technological recognition, indicative of the role of suprasegmental variables in linguistic recognition.

A framework fostered by Ng et. al (2010) was especially designed for screening rhythmic features based on mutual information. It selected 12 defining features out of 93 candidates, significantly upgrading its capability of speech recognition and fully leveraging the sensitivity of fundamental frequency (F0) to tonal languages, like mandarin. In addition, a multi-dimensional approach that combined both spatial and rhythmic paradigms of vowels was applied by Zhao and Yang (2022) to differentiate

mandarin and Fencheng dialect in Shanxi province, which revealed that rhythmic information supplemented segmental scarcity. Moreover, a comparative analysis by Li et. al (2003) on such suprasegmental features as light tone, rhotic accent, speed and accented mode, between Shanghainese mandarin and standard mandarin, concluded that rhythmic differences were the yardstick by which listeners measured how strong the accent was subjectively, especially in the face of segmental scarcity. From the perspective of machine learning, Peng (2023) fostered a dialect recognition model built on MFCC, spectrogram, among other indicators of rhythmic features. The experimental outcomes of different feature sets verified the potential of suprasegmental features for dialect recognition. The aforesaid research, mainly applicable to machine learning, provide valuable theoretical references to blaze the trail for human listeners.

In general, the disruptive research are not attitude-oriented but technology-driven. With the wide application of high-precision sophisticated kits of tech tools encompassing modelling, little do they concern the social implications entrusted by the listeners to the rhythmic features, to be more specific, their identity, image and social status embodied in intonation-represented suprasegmental features. The blank of research has gained prominence recently, as Ladegaard and Chan (2024) lay bare, in their survey on comments made by Hong Kong middle school teachers about the English accents of their students, the pivotal impact of suprasegmental features on identity formation and social classification, which echoes the outcome of research by Campbell-Kibler (2007) on South American accent that rhythm contour and phonological style led the listeners to interpret the social implications of the accent subjectively.

Two significant gaps still exist, even though previous research has gradually demonstrated the dual role of suprasegmental features in speech perception and social evaluation. First, tonal variation and perceptual mechanisms among closely related intra-dialectal varieties are not given enough attention in studies that primarily concentrate on cross-linguistic or cross-dialectal differences. Second, the two approaches are not fully integrated because attitude-oriented studies concentrate more on speech impression and identity construction, while recognition-oriented studies tend to emphasize accuracy and acoustic modeling. Therefore, future studies must cross these lines and methodically examine how suprasegmental features intersect in social categorization and dialect recognition in "micro-variation" contexts. To move dialect recognition research from a primarily technical focus toward a broader sociocognitive perspective, this study aims to investigate how listeners can both perceive subtle differences in Shenyang Mandarin and Jinzhou Mandarin intonation and assign social meaning to them in the absence of salient segmental differences.

To this end, the essay focuses on the questions as follows:

1. To what extent can listeners from different regional backgrounds perceive and differentiate between the Shenyang and Jinzhou dialects? What general patterns emerge across groups?
2. How do listeners from different regional backgrounds evaluate Shenyang and Jinzhou dialects, and in what ways do these evaluations reveal the hierarchical relations within the dialectal community?

3. Methodology

With the comprehensive application of such methods as speech collection, perception experiment, questionnaire survey and phonological analysis, the essay sets up an empirical and multi-dimensional measurement framework for research which is keenly alive to the perceptive and commentary functions of the tonal features in interrogative sentences.

3.1. Recruitment and grouping of surveyed listeners

The experiment recruits a total of 90 surveyed listeners aged around 20 on average, including 45 males and 45 females. None of the participants had formal training in linguistics, and all were undergraduate students with normal hearing. They are divided into 3 groups in accordance with their respective origins and experiences as follows:

Shenyang Group: The listeners were born and grew up in Shenyang and speaks Shenyang dialect daily.

Jinzhou Group: The listeners were born and grew up in Jinzhou and speak Jinzhou dialect daily.

Non-Local Group: The listeners originated from regions outside northeast China, and the official northeast Chinese dialect is new to them.

The grouping that draws a comparison between local and non-local language users aims to testify the consistency of the hierarchy of intra-dialectal variants in and outside their origins.

3.2. Speech recording

To secure authentic dialectal materials of which the tonal features incentivize actual feedback, the experiment invites 2 Shenyang and 2 Jinzhou dialect speakers to record the speech, all of whom are male blue-collar workers, not acquainted with Mandarin but orally representative of local dialectal features. One recognized limitation is the lack of diversity in speakers' ages and genders.

Each piece of the record is noise-free and extends 15-20 seconds. Four audio stimuli were chosen for the perception test: Shenyang speakers produced stimuli a and d, while Jinzhou speakers produced stimuli b and c.

The recording adopts the situational language method. The situation is presented by a caricature featuring a dialogue between the tourist from Southern China and the frozen pear peddler in a Northeast Chinese market. With the help of the caricature, the recorders in the role of the Northeast Chinese peddler spontaneously articulate the tonal features at the end of each interrogative question amid the dialogue, indicative of the difference in rhythmic features between Jinzhou and Shenyang dialects. All the records, processed by Praat software, keep the original tone of the recorders whose rhythm and pitch have not been edited.

3.3. Phonological comparative analysis

Praat software is applied to analyze such suprasegmental indicators, where F0 contours were the primary suprasegmental parameter that was the focus of the analysis. The goal was to verify that the intonation contours of the speech samples from Shenyang and Jinzhou were consistently different. The phonological analysis was descriptive, concentrating on F0 contours to confirm persistent distinctions between Shenyang and Jinzhou. Furthermore, inferential statistical tests were utilized in subsequent analyses of perception and attitude.

3.4. Questionnaire design

A systematic online questionnaire is set to take stock of the listeners' feedback on phonological perception and social comments. It is divided into 3 sections: linguistic background, task of phonological perception and attitude, and speech recognition task.

Section 1 collects the basic information of the listeners, including the languages they use and their familiarity with northeast Chinese dialects, which is referential to group the listeners and control the distracting variable in the subsequent analysis.

Section 2 requires the listeners to make a swift choice about whether the recorded speech they have heard is typical northeast Chinese dialect, the purity of its accent, based on an ordinal scale ranging from "very typical" to "not typical at all. Besides, the listeners shall score the recorders from 7 social aspects, including credibility, education level and sense of intimacy. The evaluative dimensions were adopted from existing language attitude scales (Zahn & Hopper, 1985), which is to explore the viability of perceptive recognition and social comment driven by suprasegmental features. The seven evaluative dimensions were later combined into a single attitude index after confirming high internal consistency (Cronbach's $\alpha \approx 0.90$).

Section 3 consists of 4 pieces of audio recordings for discrimination. The listeners are required to judge whether they belong to the same origin of the dialect and spell out their origins, if possible. Participants in the dialect comparison task answered open-ended questions by noting which dialect they thought each pair of stimuli belonged to. The section further testifies the capability of suprasegmental features for dialectal categorization.

3.5. Statistical analysis

Descriptive analyses were supplemented with inferential statistics. Binomial tests determined if discrimination accuracy surpassed random chance, whereas Fisher's exact tests evaluated group disparities in dialect naming. Chi-square tests assessed the categorization patterns of "typical Northeastern Mandarin." Wilcoxon signed-rank tests were used to compare Shenyang and Jinzhou stimuli within groups for attitude ratings. Kruskal-Wallis tests with pairwise Mann-Whitney (FDR correction) were used to look at differences between groups. Spearman correlations were employed to examine the association between dialect familiarity and evaluative preferences. We used Cronbach's α to measure how consistent the seven evaluative dimensions were with each other. All analyses were performed using SPSS (Version 27).

4. Results

The results are presented in descriptive statistics that consist of 5 parts: self-report on Praat F0 presentation, dialect familiarity, dialect perception and recognition, categorization of typical northeastern Chinese dialect, and language attitude.

4.1. Praat fundamental frequency presentation

"How much is a frozen pear? "Do you want me to peel it?"--Two typical interrogative sentences are selected from each piece of audio records for comparative analysis, presented in the form of F0 curves. F0, the objective indicator of pitch, can present the fluctuation of intonation to good effect. According to the results, as is shown in Figure 1 and Figure 2, the intonation of the Shenyang accent remains flat in general and ends up in a slight rise. Some of its segments have seen F0 indicator declines, indicative of weak interrogative feature. Figure 3 and Figure 4 showcase the intonation of Jinzhou accent fluctuates notably and

ends up in sharp rise. The Jinzhou accent, characterized by stronger interrogative feature, keeps a higher overall F0 level than its Shenyang counterpart.

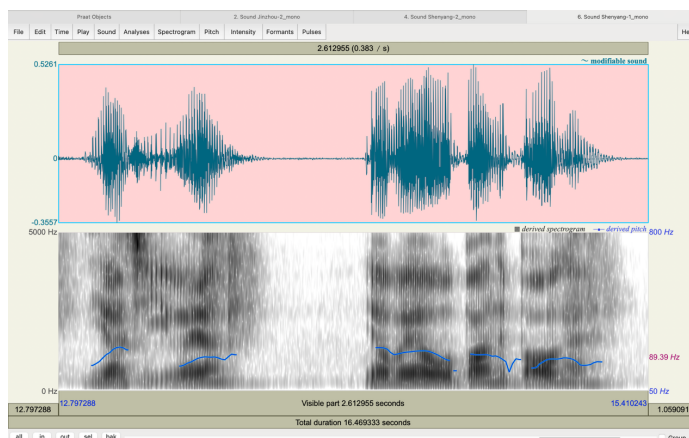


Figure 1. F0 contour of sentence in Shenyang accent (Audio A)

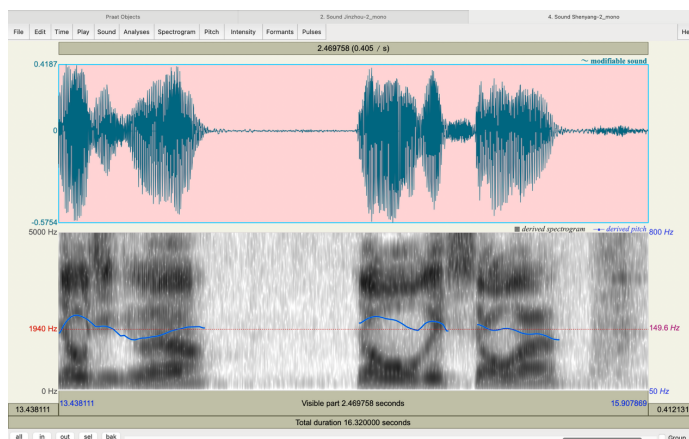


Figure 2. F0 contour of sentence in Shenyang accent (Audio D)

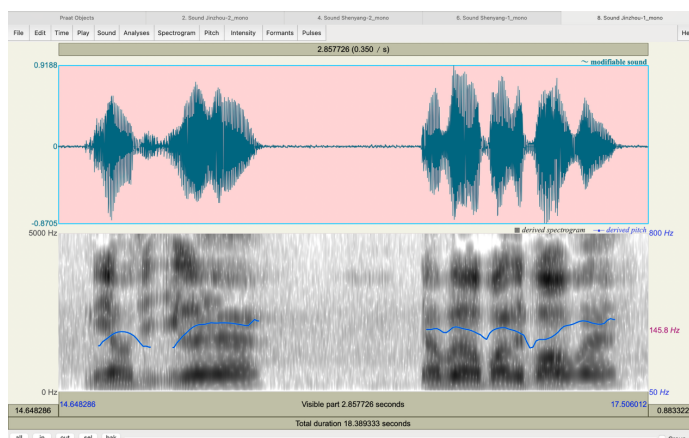


Figure 3. F0 contour of sentence in Jinzhou accent (Audio B)

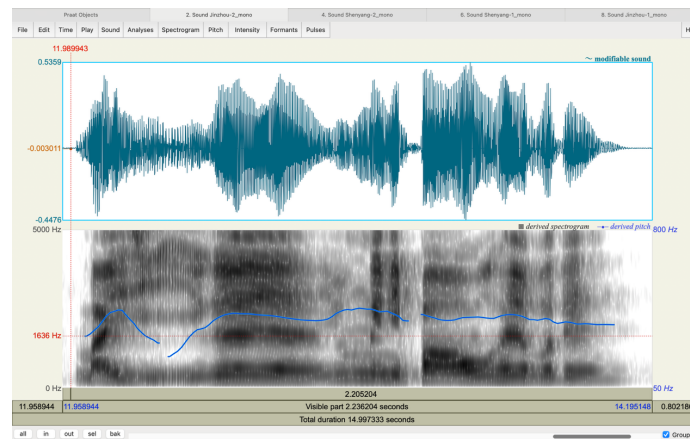


Figure 4. F0 contour of sentence in Jinzhou accent (Audio C)

4.2. Dialect familiarity

The group Jinzhou rated their own familiarity the highest ($M = 4.74$), followed by the group Shenyang ($M = 4.26$) and the non-local group ($M = 3.96$). Further analyses indicated a significant positive correlation between dialect familiarity and evaluative preference for Shenyang over Jinzhou (Spearman's $\rho = 0.27$, $p = 0.012$).

4.3. Dialect perception and recognition

4.3.1. Pair discrimination task

As is shown in Table 1, the three groups bear sharp contrast in discriminating whether the speech sounds in 6 pairs belong to the same origin of the dialect. In this regard, Group Shenyang takes the lead, as its accuracy rate of 5 pairs tops 50%; Group Jinzhou has 2 pairs above, 2 pairs at, and 2 pairs below 50%; The non-local group performs worst, with only 2 pairs topping 50% and the rest 4 lower than 50%. Moreover, only Shenyang listeners significantly distinguished same-origin pairs above chance level (63.3% correct, binomial test, $p = 0.026$, 95% CI [0.50, 0.75]). They did not do better than chance on pairs from different origins ($p = 0.26$). Neither Jinzhou (same-origin 58.3%, different-origin 47.5%) nor non-local listeners (same-origin 56.7%, different-origin 42.5%) exhibited significant performance above chance (all $p > 0.1$).

Table 1. The count of correct recognitions of the three groups in six audio comparison tasks

Group	same origin		different origin			
	Pair 3 (Shenyang A vs Shenyang D)	Pair 4 (Jinzhou B vs Jinzhou C)	Pair 1 (Shenyang A vs Jinzhou B)	Pair 2 (Shenyang A vs Jinzhou C)	Pair 5 (Jinzhou B vs Shenyang D)	Pair 6 (Jinzhou C vs Shenyang D)
Shenyang	17	21	17	14	17	16
Jinzhou	17	18	14	15	13	15
Non-local	16	18	14	12	13	12

4.3.2. Dialect origin recognition

Only a small number of listeners were able to correctly guess where the dialect came from: 3 from Group Shenyang (10%) and 2 from Group Jinzhou (6.7%). None of the listeners who were not from the area were able to do it. Binomial tests showed that all groups did much worse than chance level ($p < 0.001$). Fisher's exact test additionally indicated no significant group differences in recognition accuracy ($p > 0.1$). These results indicate that clear acknowledgment of dialect origin was significantly restricted among all listener groups.

4.3.3. Categorization of typical Northeastern Mandarin

The combined number of listeners who perceive each piece of audio record as very close or somewhat close to typical northeast Chinese dialect is collected and listed as follows.

Table 2 shows that all three groups thought that both Shenyang (a, d) and Jinzhou (b, c) recordings were typical of Northeastern Mandarin. Chi-square tests revealed no significant differences among groups for any of the four stimuli (all $p > 0.05$). Combining stimuli into Shenyang and Jinzhou also did not show any significant differences between groups (all $p > 0.1$).

Table 2. Number of participants identifying speech as typical Northeastern Mandarin

Group	Shenyang Accent		Jinzhou Accent	
	Audio a	Audio d	Audio b	Audio c
Shenyang	18	17	18	16
Jinzhou	20	13	22	19
Non-local	21	17	19	18

4.4. Language attitudes

All listeners rated Shenyang-accented stimuli (a, d) much higher than Jinzhou-accented stimuli (b, c) (Wilcoxon signed-rank test, $W = 931$, $p = 0.0015$). A Kruskal–Wallis test revealed significant group differences in the extent of this preference ($H(2) = 6.10$, $p = 0.047$). Post-hoc tests indicated that Shenyang listeners exhibited more pronounced preferences for Shenyang speech compared to Jinzhou listeners; however, this difference attained only marginal significance following FDR correction (adjusted $p \approx 0.057$). Importantly, familiarity with dialect was positively correlated with evaluative preference (Spearman's $\rho = 0.27$, $p = 0.012$), indicating that increased exposure enhances rather than reduces the bias toward Shenyang.

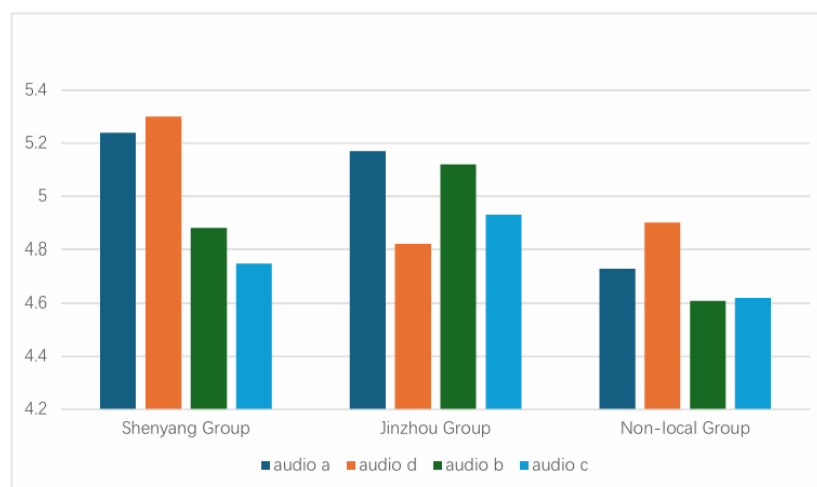


Figure 5. Mean language attitude ratings by group and audio stimulus (a–d)

Figure 5 shows that all groups rated Shenyang-accented stimuli (a, d) higher than Jinzhou-accented stimuli (b, c), and the effect was strongest among Shenyang and non-local listeners.

5. Discussion

Most listeners believe all 4 pieces of audio records to be typical of northeast Chinese dialect, but the concrete headcount of believers varies from group to group, and from piece to piece. The distributed and fluctuated pattern of recognition, the audience's failure to foster the common knowledge of what the typical northeast Chinese dialect is, irrespective of the variants they have heard. Statistical tests verified that group differences were not significant, highlighting the ambiguity of categorization. In terms of language attitude, the Shenyang dialect has received favorable comments from Group Shenyang and the non-local group collectively, whereas Jinzhou dialect fails to win consistent collective preference and has received comparatively lower rating from the non-local group. The research results have laid bare a paradox that although suprasegmental features fall short of precise recognition of dialectal identity, they shape and underpin the social status in language attitude. The non-local group's

overall lower rating could be the result of an out-group devaluation effect. They have no emotional connection to either variety, due to which their evaluative baseline is typically lower than that of local listeners (Tajfel & Turner, 1979). Empirical research on language attitudes has also confirmed similar trends, with out-groups generally evaluating unfamiliar accents less favorably (Giles & Coupland, 1991). The discussion hereby centers on the role of suprasegmental features in dialectal recognition and social comment, in terms of perceptive mechanism, attitude formation and hierarchy order.

5.1. Suprasegmental features and dialectal perception: perceptive consistency & vague labeling

In the perceptive tasks, only Shenyang listeners exceeded chance in same-origin discrimination ($p = 0.026$), while all other results were at chance level, underscoring the limited role of suprasegmental cues for explicit recognition. However, the overall accuracy rate of the perceptive tasks combined, especially the ill-performed origin recognition task, remains comparatively low, which lays bare suprasegmental features' failure to precisely categorize the origins of the dialects. As Clopper and Pisoni also consider it difficult for suprasegmental features to form categorized perception due to their consistency and volatility, the research outcome echoes their opinion but is different from the machine-based framework Ng et. al foster for screening rhythmic features as well as Zhao and Yang's multi-dimensional approach built on phonic paradigms, for the result illustrates that suprasegmental features can lead to vague sense of difference but fail to label social category in actual perceptive context. Established research mainly centers on highly efficient tech recognition or cross-dialect macro-discrepancy and loses sight of the real limit of discrimination in human perception. In this sense, the essay faithfully reflects the utmost capability of human for dialect discrimination in real life, which evolves the machine-based research into the domain of social perception.

5.2. Suprasegmental features and language attitudes: the implicit transmission of social symbolism and the attitudinal ambivalence of marginalized groups

In sharp contrast to the vagueness of perception and discrimination, language attitudes are both evident and identical, as is shown by relevant datasets. The overall preference for Shenyang was statistically significant across groups (Wilcoxon, $p = 0.0015$). There were differences between groups (Kruskal-Wallis, $p = 0.047$), with Shenyang listeners showing a stronger bias than Jinzhou listeners, but only slightly after correction. The positive correlation between familiarity and evaluative preference indicates that experience enhances rather than diminishes hierarchical attitudes. Irrespective of the non-local group's completely wrong origin recognition and low dialectal familiarity, corresponds with the viewpoint championed by Ladegaard, Chan and Campbell-Kibler that although suprasegmental features failed to trigger accurate discrimination, they carried social values subconsciously. Guided by Bourdieu's linguistic capital theory, the essay reveals that suprasegmental features, the carriers of implicit social comments, can reconstruct uneven allocation of linguistic capital without the accurate recognition of dialectal identity, which fills in the blank of micro-level research on the hierarchy backed by suprasegmental features. The discrepancy that Group Jinzhou has formed neither clear pattern of comment nor collective preference for a local dialect is of great theoretical significance. As a sidelined variant, Jinzhou dialect is worse placed to symbolic capital than its Shenyang counterpart, as Shenyang dialect grows into linguistic representative of the region justified by media circulation and sociopolitical position. Therefore, Jinzhou dialect speakers find it difficult to legitimize their local accent as a socially reputed linguistic form. Its discrepancy in linguistic evaluation lays bare linguistic insecurity proposed by Labov (1966) that users of marginalized variants may hesitate between attachment to their local dialect and awareness of its low social status. Under Bourdieu's framework, the discrepancy also signifies that symbolic power curbs the capability for self-recognition of the marginalized, so that their preference is both contradictory and vague. Therefore, the lack of a consistent evaluation pattern among Jinzhou participants highlights the unevenness of internal standardization within dialect regions, demonstrating how peripheral groups might react to internal stratification in a different way compared to central groups, and raising the possibility that this mechanism also functions in other speech communities.

5.3. Implicit approach of intra-dialectal hierarchy

The paradox between the discrepancy in categorization of typical northeast Chinese dialect and the unanimity of positive attitude toward Shenyang dialect, a showcase of Shenyang dialect's symbolic position as the intra-regional standard language, clarifies that suprasegmental features underpin linguistic hierarchy subconsciously. As Auer's interior standardization theory is based on explicit social systems and media performance and falls short of a well-grounded explanation for the paradox, the essay deepens the understanding of the implicit approach of non-standard intra-dialectal hierarchy, where implicit symbolization empowers interior standardization, even in the absence of explicit consensus on categorization. When combined, these findings suggest what could be termed a misrecognition paradox: suprasegmental cues continue to influence social assessments and perpetuate hierarchical distinctions even when listeners are unable to correctly identify dialectal origin. This paradox implies that symbolic inequality may be sustained through prosodic features without conscious awareness and draws attention to a recognition–

evaluation divide that has not been adequately addressed in earlier sociophonetic studies. Future studies on minute differences in other dialect continua and closely related language varieties may find this idea to be a helpful framework.

5.4. Limitations and expectations

Limited by insufficient material support, the essay falls short of its in-depth penetration into the psychological condition and social mindset of listeners in fulfilling the perceptive and attitude tasks. Another limitation falls on the limited representation of the samples, which may question the universal application of the conclusion, as the listeners are mainly university students. In addition, the cross-sectional design fails to capture the change in linguistic attitude with the dynamic evolution of media exposure and social contact. Lastly, the study is exploratory in scope due to its small sample size, a limited number of stimuli, and single reliance on student participants. These limitations may hinder generalizability; however, the application of inferential statistics offers preliminary yet substantial evidence for the misrecognition paradox. It was conceived as a preliminary study to explore the potential influence of suprasegmental features in intra-dialectal assessment. As a result, the findings should be interpreted cautiously, yet they offer preliminary yet compelling evidence for the suggested misrecognition paradox.

Future research may concentrate on, firstly, whether the misrecognition paradox seen here is a more universal mechanism across various sociophonetic contexts could be further investigated in future comparative studies. Besides, it is imperative to scale up the samples across different age groups and social backgrounds to verify the application of the research to various communities. Likewise, its universal application would be testified by comparative research across dialectal regions. In addition, with the application of vertical and experimental design, the incoming research would take control over media exposure and social contact, monitor the change in language attitude over time and under different circumstances, and therefore reveal the effective mechanism of suprasegmental features for social evaluation more comprehensively.

6. Conclusion

The research reveals perceptive and commentary differences in the official northeast Chinese dialect. Although the listeners have a comparatively low rate of accuracy in dialectal recognition, they have demonstrated collective preference in language attitude. To be more specific, their comments on the Shenyang dialect are more favorable than those on the Jinzhou dialect. However, Group Jinzhou fails to foster consistent collective preference, with its language attitude distributed individually. The research findings are indicative of the complicated relation between recognition and comment in dialectal perception.

Theoretically speaking, the research demonstrates and theorizes the concept of the misrecognition paradox that suprasegmental features can trigger convergent social comment, thereby reproducing symbolic hierarchy, even in the absence of accurate recognition. The phenomenon lays bare the mechanism featuring separation between recognition and comment and applies the concept of misrecognition proposed by Bourdieu to sociophonetics, signifying the capability of symbolic power for sustaining social inequality subconsciously.

Admittedly, the research has certain limitations, including a lack of diversity in age and gender among audio recorders as well as the relatively small and uneven sample sizes among groups, the exclusive use of male speakers, and the cross-sectional design, which may restrict generalizability. The limitations, to a certain extent, bear on the generalizability of research results, but does not undermine their contribution to the development of theoretical framework.

Targeted at northeast Chinese dialect as the research is, its significance goes beyond the target region. Associated dialects and language variants worldwide may have mechanisms like that of northeast Chinese dialect. In their respective context, rhythmic features can also serve as the implicit backbone that sustains the social hierarchy. In the future, research is expected to further testify the application of misrecognition paradox across languages and dialects and explore the performance and evolution of the mechanism in disparate social contexts through qualitative methods and vertical design. By fully leveraging the implicit role suprasegmental features play, the research showcases how language variants sustain social hierarchy in implicit manners.

Acknowledgements

The author declares that this research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The author further declares no conflict of interest.

References

- [1] Auer, P. (2005). Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In N. Delbecque, J. van der Auwera, & D. Geeraerts (Eds.), *Perspectives on variation: Sociolinguistic, historical, comparative* (pp. 7–42). Mouton de Gruyter.
- [2] Bourdieu, P. (1991). *Language and symbolic power*. Harvard University Press.

- [3] Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech*, 82(1), 32–64. <https://doi.org/10.1215/00031283-2007-002>
- [4] Clopper, C. G., & Pisoni, D. B. (2004). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics*, 32(1), 111–140. [https://doi.org/10.1016/S0095-4470\(03\)00009-3](https://doi.org/10.1016/S0095-4470(03)00009-3)
- [5] Giles, H., & Coupland, N. (1991). *Language: Contexts and consequences*. Thomson Brooks/Cole.
- [6] Ladegaard, H. J., & Chan, K. L. R. (2024). The importance of suprasegmental features in language attitude research: Evidence from a study of teachers' attitudes towards Hong Kong English. *Journal of Multilingual and Multicultural Development*. Advance online publication. <https://doi.org/10.1080/01434632.2024.2342923>
- [7] Li, A., Wang, X., & Yin, Z. (2003). A comparative study of Putonghua and regional varieties of Putonghua. In *Proceedings of the 6th National Conference on Modern Phonetics* (pp. 68–74). Chinese Information Processing Society of China.
- [8] Li, J., Zheng, T. F., Byrne, W., & Jurafsky, D. (2006). A dialectal Chinese speech recognition framework. *Journal of Computer Science and Technology*, 21(1), 106–115. <https://doi.org/10.1007/s11390-006-0106-9>
- [9] Lippi-Green, R. (2012). *English with an accent: Language, ideology and discrimination in the United States* (2nd ed.). Routledge.
- [10] Milroy, J. (2002). *Authority in language: Investigating standard English* (3rd ed.). Routledge.
- [11] Ng, R. W., Lee, T., Leung, C. C., Ma, B., & Li, H. (2009). Analysis and selection of prosodic features for language identification. In *2009 International Conference on Asian Language Processing* (pp. 123–128). IEEE. <https://doi.org/10.1109/IALP.2009.61>
- [12] Peng, L. (2023). *Research on Chinese dialectal language recognition based on deep learning* [Unpublished master's thesis]. Nanchang University, China. <https://cdmd.cnki.com.cn/Article/CDMD-10403-1023679253.htm>
- [13] Preston, D. R. (2013). Linguistic insecurity forty years later. *Journal of English Linguistics*, 41(4), 304–331. <https://doi.org/10.1177/0075424213502813>
- [14] Silverstein, M. (2018). Monoglot “standard” in America: Standardization and metaphors of linguistic hegemony. In D. Hymes (Ed.), *The matrix of language* (pp. 284–306). Routledge.
- [15] Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge University Press.
- [16] Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in ingroup favouritism. *European Journal of Social Psychology*, 9(2), 187–204. <https://doi.org/10.1002/ejsp.2420090207>
- [17] Wen, J. (2012). *A phonetic study of Jinzhou dialect* [Unpublished master's thesis]. Liaoning Normal University, China. <https://cdmd.cnki.com.cn/Article/CDMD-10165-1013237321.htm>
- [18] Xin, S. (2006). *A study on Sichuan university students' language attitudes toward Sichuan dialect and Putonghua* [Unpublished master's thesis]. Southwest Jiaotong University, China. <https://cdmd.cnki.com.cn/Article/CDMD-10613-2006102362.htm>
- [19] Yan, W., Tao, X., Xu, X., & Dai, Z. (2024). A statistical analysis on the functional changes of dialect and Mandarin Chinese: A diachronic perspective. *SAGE Open*, 14(2), 1–13. <https://doi.org/10.1177/21582440241255452>
- [20] Zahn, C. J., & Hopper, R. (1985). Measuring language attitudes: The speech evaluation instrument. *Journal of Language and Social Psychology*, 4(2), 113–123. <https://doi.org/10.1177/0261927X8500400203>
- [21] Zhang, Y. (2014). A comparative study of the phonetics and vocabulary of Shenyang dialect and Putonghua. *Jintian*, 3, 301–305. <https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&filename=JINT201403012>
- [22] Zhao, Q., & Yang, J. (2022). A study on speaker identification between Chinese dialects and Mandarin based on vowel acoustic space. *Chinese Forensic Science*, 6, 67–74. <https://doi.org/10.3969/j.issn.1671-2072.2022.06.009>
- [23] Zheng, J. (2007). *The influence of dialect experience on the perception of Mandarin phonology: Evidence from Minnan dialect bilinguals* [Unpublished master's thesis]. Fujian Normal University, China. <https://cdmd.cnki.com.cn/Article/CDMD-10394-2007098746.htm>
- [24] Zhou, M. (2001). The spread of Putonghua and language attitude changes in Shanghai and Guangzhou, China. *Journal of Asian Pacific Communication*, 11(2), 231–253.