

# A Sociophonetic Investigation of Chinese Gay Couples' Variability of Pitch Properties in Vlogs

Xinyu Liao

**Abstract**—Despite the accumulating body of research in sociophonetics exploring gay men's pitch characteristics (i.e., mean vocal pitch and pitch range), previous studies usually investigate a uniform concept of 'gay men's speech' by comparing heterosexual and gay men's pitch properties. However, results were contesting and inconsistent across various studies regarding the pitch properties (pitch ranges or mean voice pitch) of gay men. Instead of treating gay men's speech as a unified speaking style, this paper investigates the multiplicity of gay speaking styles by exploring the intra-group pitch variations among 20 pairs of Chinese gay couples in their self-shot videos. Specifically, the present study compares the pitch properties, including the mean vocal pitch, pitch range, and pitch variability, between those Chinese gay men who self-position as 'lao gong' (husband) and those who self-identify as 'lao po' (wife) in their love vlogs (video blogs). These videos normally last from 5 to 10 minutes on a Chinese online video sharing platform - 'Bilibili.' After dividing these gay couples' utterances into intonational phrases, I used the speech analysis software named Praat to measure the average pitch, pitch range (the maximum pitch value minus minimum pitch value), and pitch variability (the standard deviation of pitch values) on each intonational phrase. Compared with those 'gay husbands,' results showed that those 'gay wives' would speak with higher-pitched voices ( $p < 0.05$ ), wider pitch ranges ( $p < 0.0001$ ), and more variable pitch values ( $p < 0.0001$ ). When locating the discourse functions of these pitch characteristics in their vlogs, I argue that those 'gay wives' frequently utilize the so-called 'pitch dynamism' to construct an expressive and cute 'wife' persona in intimate discourse.

**Index Terms**—Gay men's speech, Mandarin Chinese, pitch, sociophonetics.

## I. INTRODUCTION

Interlocutors' utterances and voices transmit the referential message of the content and provide rich information about speakers' self-identification, such as gender, education level, and ethnicity [1]-[3]. It might be oversimplified to assume a direct association between one linguistic form and a social category because of the speakers' agency and the individual variability. However, it is still meaningful to investigate the overall patterns across various identity categories since previous studies have shown statistically significant differences between individuals of different sex due to gender ideologies [1].

Interest in investigating the influence of sexual orientation on speech production can be first traced to Gaudio's experiment in the 90s of the last century [4]. Against people's general stereotypes about gay men's feminine speaking styles

(with high pitch) as deviant from the heteronormative voice, this experiment did not find any significant differences between gay men and heterosexual men's average pitch (measured by the acoustic correlate – fundamental frequency or F0). Concurring this finding, many studies in the following years found no significant differences between heterosexual and gay men's average voice fundamental frequency [5]-[7].

A recent study by Geng and Gu in China even reported a striking result that Chinese gay men speak with a lower-pitched voice than their heterosexual counterparts, and they explained it might be due to gay subjects' unwillingness to conceal the gay identity [8]. Although previous studies provide fruitful results on the speaking styles in the gay communities, their research is mostly based on a homogeneous concept of 'gay men's speech' compared to heterosexual males' in experimental settings. This paper then further explores the potential multiplicity within gay men's community by investigating the intra-group pitch variations among 20 Chinese gay couples in their self-shot vlogs (video blogs) posted on an influential Chinese video-sharing platform - Bilibili.

## II. THEORETICAL FRAMEWORKS

As mentioned before, although it might be tempting to label one linguistic form (e.g., a grammatical or a phonetic form) as a feature of a particular identity category, one form can have multiple functions in various contexts. The theory of indexicality can be used here to approach the relationship between sexuality and pitch properties in the current study [9], [10]. From Ochs's perspective, the relationship between gender identity and language at various levels (e.g., morphology, syntax, phonology, and discourse) is usually not in a direct manner [9]. Direct mappings between gender and linguistic forms are, for instance, the third person pronouns (he/she) and some honorifics (sir/madame) in English. Besides these exceptions, Ochs argues that social acts and stances mediate the connection between a linguistic form and gender identities [9].

As a central component of Ochs' theory, stance here or stancetaking can be conceptualized as how speakers would position themselves regarding their relationships with the referential meaning of the utterances, other speakers, identity categories, and ideologies [11]. In this sense, the stances or social acts locally indexed by linguistic forms can further construct macro-social categories (e.g., gender and ethnicity) under certain culturally anchored ideologies.

Under Silverstein's ideological model of indexicality, this process begins with an 'n-th indexical order' [10]. The n-th order indicates that links between linguistic forms and social

Manuscript received October 27, 2021; revised May 5, 2022.

The author is with University of Nottingham Ningbo China, China (e-mail: hnyx15@nottingham.edu.cn).

group memberships are gained through their correlations. Furthermore, the  $n$ -th order indexical order can always be reconstructed and reinterpreted by speakers' agency to form a new ' $n+1$ st indexical order'. Silverstein uses Labov's work in New York to explain the model [10], [12]. In the ' $n$ -th indexical order,' speakers' frequent usage of standard pronunciations in Labov's study can correlate with the membership of higher social class. This correlation could further reconstruct the meaning of 'prestige' ( $n+1$ st indexical order) for speakers within the ideological space of ' $n$ -th indexical order.' Following Ochs and Silverstein's indexicality approach to language and identity, this paper will also locate gay men's pitch variations at interactional contexts to identify the stance,  $n$ -th, and  $n+1$ st indexical orders associated with pitch properties under the gender ideologies in China. These can then help to decode the socio-indexical factors underlying pitch variability.

### III. METHODOLOGY

For investigating the intra-group variations among Chinese gay men, this study compiles a spoken corpus by randomly collecting 30 love vlogs (video blogs) made by 30 pairs of gay couples in a Chinese video-sharing platform – Bilibili.

The topics in those love vlogs are about these gay couples' dating experiences, how they met each other, and some conversations about their daily lives. The duration of these love vlogs usually lasts between 5 and 10 minutes. This data collection differs from most of the previous studies comparing gay men and heterosexual reading speech in experimental settings [5]-[7]. In these vlogs, their gay identity might thus be more salient when these couples interact with each other. Interestingly, when qualitatively analyzing the main content and themes in these videos, 20 pairs of gay couples seemingly performed heterosexual 'husband and wife' relationships in their vlogs. Among these couples, one usually self-positioned as 'lao gong' (husband), and the other usually self-identified as 'lao po' (wife) in their same-sex intimate relationships. Therefore, this study uses 'gay husbands' and 'gay wives' to label these two local identity categories. This paper would only focus on these 20 pairs of gay couples to explore how these gay men with divergent self-positionings will use their speech to construct their identities.

When analyzing the pitch properties among these gay couples, this paper firstly divided their utterances into intonational phrases and annotated them on PRAAT [13]. Boundaries of each intonational phrase were marked following three criteria – pauses, lengthening of the final syllables, and resetting of the pitch level [14]. After annotating the intonational phrases and the corresponding speaker types (gay husbands/gay wives), this study used PRAAT to acoustically measure the mean fundamental frequency (F0), F0 range (maximum F0 minus the minimum F0), and F0 variability (standard deviation of F0) on each intonational phrase. Although F0 does not equal pitch, F0 is the acoustic correlate of the pitch in the phonetic analysis [15]. Hence, the average F0, F0 range, and F0 variability correlate with the average pitch, pitch range, and pitch variability, respectively.

After data collection of the pitch properties among 20 gay

couples, this paper used independent-samples  $t$ -tests to compare the pitch variations between two groups (gay husbands and gay wives). After quantitative analysis of pitch variations, discourse functions performed by pitch properties will also be qualitatively examined in contexts. Furthermore, this study will link such discourse functions with the indexicality (stancetaking,  $n$ -th, and  $n+1$ st indexical orders) of the pitch properties.

By combining both quantitative and qualitative methods, this study will attempt to uncover both the variation patterns of pitch characteristics and the socio-indexical factors underlying these variations.

### IV. PITCH VARIATIONS AMONG GAY COUPLES

This study firstly runs statistical analysis on the pitch variations between two local identity groups. This paper uses three boxplots (Fig. 1 – Fig. 3) to present the full stretch of the data on the average vocal pitch, pitch range, and pitch variability.

Regarding the average pitch between gay husbands and gay wives, Table I and Fig. 1 show that gay wives speak with higher-pitched voices than husbands ( $p=0.02<0.05$ ). As shown in Fig. 1, some of the gay wives' average F0 has even been above 160 Hz and one speaker's mean F0 even reaches 206 Hz in his love vlog. These pitch values are remarkably higher than Mandarin-speaking males' average pitch (150 Hz) [16]. Therefore, gay wives might appropriate the use of higher vocal pitch to perform their more feminine persona as 'wife' in the vlogs when they talked about their intimate relationships. Moreover, as seen in Fig. 2 and Table I, gay wives also used significantly wider pitch ranges ( $p<0.0001$ ) when talking with their counterparts in their love vlogs. As shown in Fig. 2, most gay husbands' pitch ranges center around 75 and 85 Hz, whereas gay wives' pitch ranges cluster around 120 Hz and 155 Hz. Concurring with this result, the results on pitch variability also shows a similar trend in Table I and Fig. 3. Gay wives adopted more variable and dynamic pitch values ( $p<0.0001$ ) when they spoke with gay husbands in the vlogs. Consistent findings on pitch ranges and pitch variability between gay husbands and wives point out that gay wives would speak more 'dynamically' and 'dramatically' when they constantly changed and fluctuated their vocal pitch across their utterances in the vlogs. In contrast, most of the gay husbands' mean vocal pitch remained within the scope of a normal mandarin male speaker's mean pitch (150 Hz) [16].

Although most of the previous studies cited above did not find significant differences between gay and heterosexual males' pitch properties, this study finds some intra-group variations within the gay community. These findings on pitch variations thus suggest that when those gay wives consciously adopted a 'wife' persona in their self-presentations online, they would reappropriate the phonological resources associated with stereotypical feminine speech styles [17]. In contrast, when performing 'husband and wife' relationships online, gay husbands made the 'unmarked' choice of using a speech style that is more congruent with a male speaker compared with their counterparts. In other words, when presenting as a 'husband' in a same-sex relationship, gay men would stick to a more stereotypical male's voice with low pitch, narrower pitch

ranges, and less variable pitch values.

TABLE I: MEANS FOR EACH VARIABLE &amp; T-TEST RESULTS

Role types	Mean F0 (Hz)	F0 range (Hz)	F0 variability (Hz)
Gay husbands	124.81	79.34	22.96
Gay wives	139.81	134.37	36.06
	$P < 0.05 *$	$P < 0.0001 ***$	$P < 0.0001 ***$

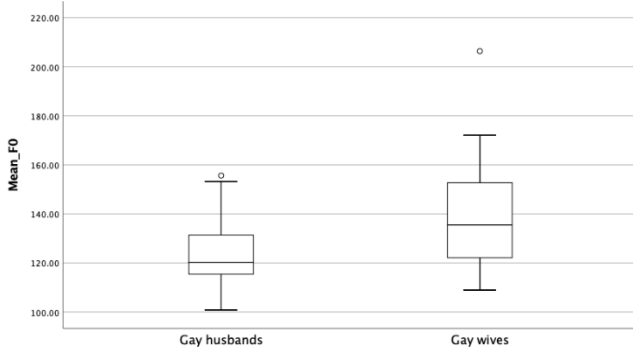


Fig. 1. Mean F0 (Hz) by group.

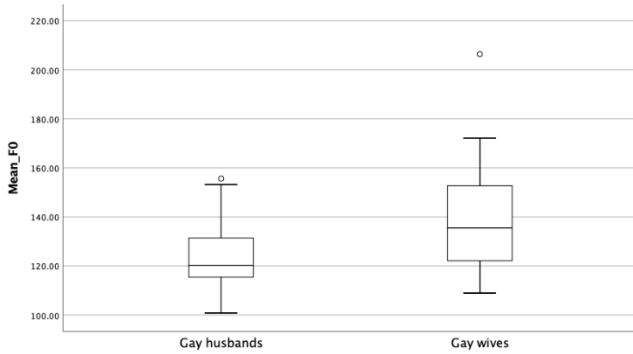


Fig. 2. F0 range (Hz) by group.

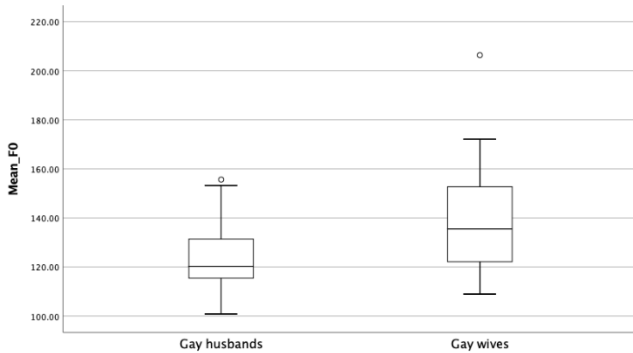


Fig. 3. F0 variability (Hz) by group.

## V. INDEXICALITY OF THE PITCH PROPERTIES

Based on the quantitative results reported in the last section, this section will qualitatively examine the pitch patterns within the conversations to understand the socio-indexical and semiotic underpinnings of the pitch variations. When closely investigating the intimate discourse between gay wives and gay husbands, this study finds that the socio-pragmatic or discourse functions of the high-pitched voice with wider pitch ranges and more variable pitch values are expressive and emotive. Excerpt 1 (with English translations

under each line) is from a short conversation between a gay husband (Dong) and a gay wife (Lee) in their love vlog. This interaction happened when Lee and Dong hung out in a city and casually chatted with each other. In this conversation, Dong's mean pitch was approximately 104 Hz, and his pitch range was about 45 Hz. In contrast, Lee's maximum vocal pitch reached 467.61 Hz, and his pitch range was over 150 Hz. Dong used a joke to tell Lee that he loved Lee by using Chinese words (ai lao hu you) with similar pronunciations to 'I love you' in English. This joke elicited Lee's affective response with a cute 'sajiao' persona [18] to 'complain' about Dong's joke.

Excerpt 1:

1. Dong: Deng xia qu mai ping you ba ↑  
Let us buy a pot of oil later ↑
2. Lee: Shen me you?  
What kind of oil?
3. Dong: ai lao hu you (@ @ @)  
ai lao hu you (I love you) (@ @ @)
4. Lee: Oh (2) ni hao e xin a :::  
Oh (2) you are so cheesy :::
5. ni shi zhen de you (.) wo he ni JIANG  
you are greasy (.) I must TELL YOU
6. [tai you ni la ↑  
[So greasy ↑
7. Dong: [Zeng me la?  
[Why?
8. Lee : wo hai yi wei ni yao shuo se la you ne  
I initially thought that you meant the salad oil

In Chinese culture, the 'sajiao' persona is associated with childish behaviors and usually links to girls' intimate affect performance with their friends or lovers [19]. The conversational data in this study and the example above find that the interaction between 'sajiao' and the intimate affect performance of 'girlfriend' personae is not constrained to females. For instance, Lee also performs a 'sajiao' persona by repeating his 'complaints' about Dong's joke with a higher pitch and wider pitch ranges in this conversational context. In this way, he created an image of a cute 'wife' who was shy and unwilling to receive his husband's overtly romantic expressions in public. Furthermore, although the specific speech acts (such as 'complaining' in this example) performed by a high pitch and more dynamic pitch patterns vary across contexts, the primary or overarching stances indexed by these pitch properties are emotive and expressive.

As mentioned before, stance or stancetaking relates to how interlocutors would evaluate their content of utterances, social relationships with others, and memberships with macro social categories. Regarding higher pitch and dynamic pitch patterns, the expressive or emotive stances indicate that gay wives more frequently utilize these pitch characteristics to take more affective-oriented stances towards their utterances in their intimate relationships with the 'gay husbands' counterparts. These stances then ideologically index or link with the more feminine 'gay wife' persona they have self-positioned in their love vlogs. In contrast, gay husbands' 'unmarked' choice of speaking with low pitch and less fluctuating pitch patterns then index their opposing stance towards more affective-oriented acts. These opposing stances

can also be intertwined with the ideological concept of hegemonic masculine ‘husband’ built upon the disengagement of expressiveness and flamboyance in Chinese culture.

This paper will also use Silverstein’s model of indexical orders to have a more nuanced analysis of the ideological processes behind the indexicality and decode the relationships between the pitch properties, local group identities, and macro social identities [10], [20]. As shown in Table II, to attain the socio-indexical meanings of construing ‘husband and wife’ relationships, the differentiation between higher-pitched voices with dynamic pitch patterns and low-pitched voices with stable pitch patterns should undergo a transition from (*n*-th order) gender-based variation to (*n*+1st order) distinction of role types in same-sex intimate relationships.

TABLE II: INDEXICAL ORDERS OF PITCH PROPERTIES

Indexical orders	Higher-pitched voice & more dynamic pitch patterns (wider pitch ranges & more variable pitch values)	Lower-pitched voice & more stable pitch patterns (narrower pitch ranges & less variable pitch values)
<i>n</i> -th order: gender-based variation	Females’ stereotypical speech styles	Males’ stereotypical speech styles
<i>n</i> +1st order: distinction between the role types (husband vs. wife)	‘Gay wives’: cute, like doing ‘sajiao’ (childish acting), more expressive and flamboyant	‘Gay husbands’: more rational and calmer, always pampering wives (inclusive towards wives’ demands)

In this process of shifting from the *n*-th indexical order to the *n*+1st indexical order, gay wives and gay husbands reappropriate and transcend the indexical meanings of gender variations to construe distinctions of their roles in the relationships. In this sense, gay wives’ use of high pitch and more fluctuating pitch patterns index the expressive and emotive stances and ideologically link the stereotyped image of girlfriend personae behind such pitch characteristics to the performance of the gay wife persona. Based on the opposing stances towards affective acts as mentioned before, gay husbands’ use of low pitch and more stable pitch patterns can also associate the image of heterosexual boyfriend personae with their positioning as ‘gay husband,’ since inexpressiveness constitutes a core of hegemonic masculinity in China [21].

From the analysis above, the indexicality of pitch variation patterns between gay husbands and wives can be realized through the differentiation in stancetaking – expressive or inexpressive. These stances further help gay couples transcend the macro gender-based distinction towards local interpersonal distinctions of role types (husband and wife) in intimate relationships.

## VI. CONCLUSION

This paper sets out to investigate the multiplicity of gay men’s speech by exploring the pitch variation patterns within gay couples in their self-made love vlogs posted online. In

their conversations, this study finds that in contrast to gay men with self-positioning of being ‘lao gong’ (husbands), gay men with self-positioning of being ‘lao po’ (wives) used higher pitch, wider pitch ranges, and higher variability (standard deviation) of pitch values in each intonational phrases. These pitch characteristics can be used to index certain stances (expressive or inexpressive) and to reappropriate the social meanings derived from (*n*-th indexical order) existing gender-based phonetic variations to construe the (*n*+1st indexical order) local ‘husband and wife’ relationships. This finding also, to some extent, challenges the conceptualization of gay men’s speech as a homogenous concept in contrast to heterosexual men’s speech in some of the previous studies [5]-[9]. Furthermore, variationist sociolinguistic studies about Chinese gay men’s speech in the future can include other variables, which have gendered meanings such as the fronting levels of fricative sibilants- /s/ and /ʃ/ in Mandarin Chinese [22], [23]. Moreover, the multimodal resources accompanying the speech, such as gestures, facial expressions, and dressing styles, are also worth further exploration.

## APPENDIX

Transcription conventions for the excerpt in analysis [24]  
(xxxx) Parentheses with xx indicate talk which is undecipherable for the transcriber

((Water)) Words in double parenthesis indicate noise from an external sound

(@@@) Laughter: The length of the @ indicates the length of the laughter

.hhh The letter h indicate noticeable aspiration, its length roughly proportional to the length of the aspiration. If preceded by a dot it indicates in-breath, if followed by a dot it indicates out-breath

[ Left sided bracket indicate where overlapping speech occurs

° Indicates talk which is noticeably quieter than the surrounding talk

> < ‘Greater than’ and ‘less than’ symbols indicate talk which is noticeably faster than the surrounding talk

(parents) Words in parentheses indicate transcriber’s comments not transcriptions

(1.5) Numbers in parentheses indicate the periods of silence in tenths of second

(.) Indicates a pause in talk of less than 0.2 seconds

Becas- A hyphen indicates words which are incomplete because of abrupt cut off or self-interruption

He says Underlined words indicate stress or emphasis

↑ An arrow pointing upwards indicate a marked pitch rise

= Equal signs indicate latching with no noticeable silence between the talk of different people

::: Colons indicate the sound was prolonged

WHAT Capital letter indicate an increased volume in talk

I[...] Parentheses with three dots indicate that there is a gap between the sections of the transcription which were not included

**Think** Words in bold indicate phrases and words which the transcriber wishes to highlight

# CONFLICT OF INTEREST

The author declares no conflict of interest.

# AUTHOR CONTRIBUTIONS

Xinyu Liao conducted the research, performed the data analysis, and wrote the paper. He had approved the final version.

# ACKNOWLEDGMENT

I would like to express my gratitude for the guidance from Dr. Ran over doing phonetic analysis with Praat. Moreover, I also want to express many thanks for the feedback and suggestions from the reviewer.

# REFERENCES

- [1] C. T. Ferrand and L. B. Ronald, "Gender differences in children's intonational patterns," *Journal of Voice*, vol. 10, no. 3, pp. 284-291, 1996.
- [2] Y. Zhang. (July 2014). A corpus-based analysis of lexical richness of Beijing Mandarin speakers: Variable identification and model construction. *Language Sciences*. [Online]. 44. pp. 60-69. Available: <https://doi.org/10.1016/j.langsci.2013.12.003>
- [3] A. Szakay. (June 2012). Voice quality as a marker of ethnicity in New Zealand: From acoustics to perception. *Journal of Sociolinguistics*. [Online]. 16(3). pp. 382-397. Available: <https://doi.org/10.1111/j.1467-9841.2012.00537.x>
- [4] R. P. Gaudio, "Sounding gay: Pitch properties in the speech of gay and straight men," *American Speech*, vol. 69, no. 1, pp. 30-57, 1994.
- [5] S. E. Linville, "Acoustic correlates of perceived versus actual sexual orientation in men's speech," *Folia Phoniatrica et Logopaedica*, vol. 50, pp. 35-48, 1998.
- [6] R. Smyth, G. Jacobs, and H. Rogers, "Male voices and perceived sexual orientation: An experimental and theoretical approach," *Language in Society*, vol. 32, no. 3, pp. 329-350, 2003.
- [7] B. Munson, "The acoustic correlates of perceived masculinity, perceived femininity, and perceived sexual orientation," *Language and Speech*, vol. 50, no. 1, pp. 125-142, 2007.
- [8] P. Geng and W. Gu. (March 2021). Acoustic and perceptual characteristics of Mandarin speech in gay and heterosexual males. *Language and Speech*. [Online]. pp. 1-14. Available: <https://doi.org/10.1177/00238309211000783>
- [9] E. Ochs, "Indexing gender," in *Rethinking Context*, D. Alessandro, G. Charles, Ed. Cambridge, U.K.: Cambridge University Press, 1992, ch. 14, pp. 335-359.
- [10] M. Silverstein, "Indexical order and the dialectics of sociolinguistic life," *Language and Communication*, vol. 23, pp. 193-220, 2003.
- [11] J. W. D. Bois, "The stance triangle," in *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*, R. Englebreton, Ed. Amsterdam, The Netherlands: John Benjamins, 2007, ch. 6, pp. 139-182.
- [12] W. Labov, *Sociolinguistic Patterns*, Philadelphia, Pennsylvania: University of Pennsylvania Press, 1972.
- [13] P. Boersma and D. Weenink. (8 September, 2021). Praat: Doing phonetics by computer [Computer program]. Version 6.1.53. [Online]. Available: <http://www.praat.org/>
- [14] A. Cruttenden, *Intonation*, Cambridge: Cambridge University Press, 1986.
- [15] L. Zimman. (June 2018). Transgender voices: Insights on identity, embodiment, and the gender of the voice. *Language and Linguistics Compass*. [Online]. 12(8). pp. 1-16. Available: <https://doi.org/10.1111/lnc3.12284>
- [16] K. M. C. Marjorie, "Gender differences in the Chinese language: A preliminary report," in *Proc. the 9th North American Conference on Chinese Linguistics (NACCL)*, vol. 2, 1998, pp. 35-52.
- [17] E. Levon, "Sexuality in context: Variation and the sociolinguistic perception of identity," *Language in Society*, vol. 36, no. 4, pp. 533-554, 2007.
- [18] R. L. Starr, T. Wang, and C. Go, "Sexuality vs. sensuality: The multimodal construction of affective stance in Chinese ASMR performances," *Journal of Sociolinguistics*, vol. 24. pp. 492-513, February 2020.
- [19] K. Hardeman, "Gender and second language style: American learner perceptions of use of Mandarin saijiao," Ph.D. dissertation, Dept. Linguistics, University of Hawaii., Honolulu, HI, 2013.
- [20] Q. Zhang, *Language and Social Change in China: Undoing Commonness through Cosmopolitan Mandarin*, 1st ed. New York: Routledge, 2018.
- [21] K. Louie, *Theorising Chinese Masculinity: Society and Gender in China*, Cambridge: Cambridge University Press, 2002.
- [22] X. Liao, "A sociophonetic study of identity co-construction among Chinese gay couples in vlogs," presented at the 3rd International Conference on Languages, Linguistics and Society, Universiti Malaysia Sabah, September 14-15, 2021.
- [23] F. Li. (September 2017). The development of gender-specific patterns in the production of voiceless sibilant fricatives in Mandarin Chinese. *Linguistics*. [Online]. 55(5). pp. 1021-1044. Available: <https://doi.org/10.1515/ling-2017-0019>
- [24] J. Potter, "Discourse Analysis as a way of analysing naturally occurring talk," in *Qualitative Research: Theory, Methods and Practice*, S. David, Ed. London: SAGE Publications. 2004, ch. 5, pp. 200-221.

Copyright © 2022 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

**Xinyu Liao** was born in 2000 in Changsha, Hunan province. He is an undergraduate student at the School of Education and English, University of Nottingham Ningbo China. He is interested in sociolinguistic variation on language and gender by using phonetic analysis, corpus-assisted lexical-grammatical analysis, and psycholinguistic techniques (e.g., reaction time, eye-tracking, and ERPs).