Can L2 Speakers Perceive Sociopragmatic Nuances at the Sound Level? Exploring Chinese Listeners' Perceptions of (ING)

Ke Lin

Georgetown University
kl863@georgetown.edu

ABSTRACT

The variation between the alveolar and velar forms (ING) has been widely used in conveying social meanings for native speakers in many English varieties. L2 speakers' perception of (ING), however, remains unknown. The presented study employs a matched guise task exploring Chinese listeners' mental representations of the (ING) variable, and a social network questionnaire, assessing how their associations interact with the levels of integration in the host society. Results from multivariate analysis show that (1) listeners with higher social network strengths perceived the speakers to be less educated, regardless of the (ING) variant; (2) the velar variant was perceived as more educated; (3) speakers' gender, and regional background had greater impacts on listeners' ratings than the (ING) guise.

Keywords: sociophonetics variation, perception, L2 processing

1. INTRODUCTION

We conduct a perception study on L2 speakers' perceptibility and judgments of a well-documented, stable phonological English variable (ING). For native speakers, the most common realizations of (ING) is the standard velar form, *-ing*, and the non-standard, alveolar form, *-in.* A vast amount of research has documented that in addition to the standard and non-standard differences, the variations have been used consistently in conveying social meanings in many English varieties [1,2,4,5]. Cross-dialectically, the standard velar variant is associated with prestige and formality while the alveolar variant with lack of sophistication and articulateness. While variations in productions and perceptions of (ING) among native speakers received much attention, no consistent patterns or clear explanations have been given to the variation of (ING) for immigrants in Anglophone countries. To bridge the gap, we investigate the perception of (ING) by native speakers of Chinese living in the U.S. We seek to uncover the following meanings that L2 speakers ascribe to the variants of (ING):

i. How salient is (ING) in speech to Chinese speakers in the U.S.?
ii. What are the meanings of (ING) for L2 speakers in the U.S. and to what extent do associations between (ING) variants and the social meanings resemble those of the native speakers, as demonstrated in Campbell-Kibler [5] and Kiesling [4]?
iii. Is hearers' degree of integration into the L1 society related to their evaluation of (ING)?

We envision results of this study may inform e.g., language and communication training. By understanding how L2s perceive and interpret the sociolinguistic cues associated with phonological variants, educators can improve lessons to support learners in developing sociolinguistic competence and navigating social meanings. By emphasizing the influence of social networks on perception, we also highlight the importance of intercultural communication skills for successful social integration.

2. BACKGROUND

2.1 Meanings of (ING) for L1 speakers

Early studies found that both in the U.S. [1] and the U.K. [3], men used the informal alveolar variant *-in* more than women. The alveolar variant was used more in Southern U.S. states [2], and the velar variant was used more often in formal speech. Newer studies uncovered more nuanced social meanings that (ING) contribute to the speakers' identity and stance. Kiesling [4] observed that boys in fraternity groups used the informal variant *-in* to index *solidarity, roughness, masculinity*, and *-ing* when exhibiting structural *power*. Thus, the social meanings of the variants evoke a wide range of impressions depending on the context. At the perception level, impressions of the (ING) variants

are even more complex. Campbell-Kibler [5] performed a matched guise study in which she matched variants with both Southern and Californian speakers' speech. The alveolar *-in* variant strengthened perceptions of the Southern accent, *informal*, *masculine, country, laidback,* and *lack of education,* the velar *-ing* evoked perceptions of *formal, urban, gay,* and *metrosexual*.

2.2 (ING) production by L2 speakers

Considering L2 English speakers, Schleef [6] found that Polish teenagers who interacted mainly with Scottish or mixed Polish-Scottish teens preferred to use -in more often than Polish teens who interacted only with other Polish teens. Drummond [7] observed that adult Polish women produced the alveolar form more frequently than men, which is the opposite pattern observed in native speakers. They speculated that women worked jobs where it was easier to adopt the local vernacular variant. However, this explanation does not apply to the teenagers in Schleef's study, as both boys and girls had similar exposure to native speech. Therefore, other factors must contribute to the difference in usage between gender and local variants. While the velar form is associated with sophistication and formality for L1s, L2 speakers (of varied proficiency and contact level with L1s) may not share the same understanding. As such, L2 speakers' backgrounds and the types of interactions with L1 speakers may contribute to a different encoding process of the (ING) forms. Expanding on how social clusters shape linguistic practice (e.g., as demonstrated in Milroy [8,9]), we include L2 speakers' levels of interaction with native speakers in the analysis.

3. METHOD

We conducted a matched guise study combined with a perception questionnaire to investigate L2 speakers' interpretations of social meanings associated with variants of (ING). Our study largely followed Campbell-Kibler's [5] procedure for consistency with previous work, but we used different strategies for speech elicitation and generating audio. In addition, we included a survey on language use and a social network to assess participants' integration in academic or professional life, living arrangement, and social and recreational life based on Ager & Strang [10]. While the matched guise task captured participants' covert associations with (ING) variants, the language use and social network survey explored the social conditions influencing L2s' perceptions of the variants.

3.1 Participants

Study participants were recruited from various parts of the U.S. as individuals who identified any dialect of Chinese as their first L1 language, and who have lived in the U.S. for at least one year. The minimum length of stay requirement ensured that they had first-hand exposure to local native speakers in their communities, and enough time to develop their local networks. 89 participants at ages ranging from 18 to 55 took part in this study virtually from home.

3.2 Audio stimuli

Audio recording comes from two native speakers of American English from California and two from the South (Georgia, Mississippi), collected over Zoom. First, each speaker was asked to tell a short story based on a series of comic strips [11] using as many present progressive verb forms as possible. The speakers then re-recorded segments with (ING) tokens in both alveolar and velar forms, matching the original intonation and speed. Using the velar and alveolar (ING) variants of the segments, we used the *cut* and *concatenate* functions in Praat [12] to replace the (ING) words with the reproduced alveolar and velar counterparts, yielding two pairs of manipulated recordings for each original story. We selected two 20- to 30-second-long segments from each speaker in each guise, yielding 8 audio clips, each containing 4 to 8 tokens of (ING) words. Every pair of recordings matches in every aspect except for the production of (ING) words. To reduce artifacts due to the manipulation, we manually adjusted the volume, pitch, and speed of each inserted segment to match the surroundings.

Qualities	*arrogant, irritable, hardworking, effective communication skills, perfunctory, confident, good management skills, logical/organized, intelligent, well-educated, has a high salary, humorous, lazy, not fit*
Personas	*student, boss, stoner, jock, urban dweller*
Stereotypes	*redneck, stoner, blue-collar*

Table 1: Social impressions elicited from Chinese listeners selected into the survey instrument.

3.3 Pilot interview

We first played the eight manipulated clips to five Chinese listeners and elicited opinions on the audio quality and speakers. We then incorporated the most frequently reported impressions into the experiment materials (see Table 1 for the selected impressions).

3.4 Experiment materials

3.4.1 Matched guise task survey

Next, we exposed each participant to 8 different parts of the 8 audio clips, each with a fixed *speaker guise:* 4 speakers with the *-ing* form, and 4 using *-in*. Participants evaluated each speaker based on (1) the levels of their 7 perceived personal *traits* on a continuous 6.0 Likert scale (*formal, educated, warm, extroverted, intelligent, engaging, talks very fast*), (2) the presence of social attributes from Table 1 (*qualities*, *personas*, and *stereotypes)*, as yes/no/skip answers, and (3) gender (*male, female*), ethnicity (*White, Black, Asian, Hispanic or Latino, unsure*), region (*New England, South, East Coast, West Coast, Other, unsure)*. The presented audio parts were selected by splitting the 8 clips in halves, re-assembling in two survey versions A, B ([14,15]), and assigning to participants at random such that each would hear all original 4 speakers with both -ing and -in in their survey.

3.4.2 Language and social network questionnaire

In the questionnaire, we collected participants' anonymized *id*, *age, age of arrival in the U.S.,* and *gender*. We characterized their engagement with the host community by surveying six social networks: *workplace, collaborators, interest groups, friends, previous housing, and current housing*. Participants assessed each network based on if it consisted of *mostly Americans (+1), mostly Chinese (+0), balanced Chinese and Americans (+0.5),* or *others (+0)*. We assigned scores in the brackets and aggregated them across all 6 networks to obtain the overall *social network strength* score, between 0 (lowest) and 6 (highest strength).

4. DATA ANALYSIS

Surveys were delivered via Qualtrics [15]. After eliminating incomplete and invalid responses (e.g., submitted in under 5 minutes), we obtained 41 responses for Survey A [13], and 42 responses for Survey B [14]. We applied multivariate analysis using Rbrul [16] to infer how various aspects of speakers and survey participants interacted with participants' social perceptions of the speakers. We included seven potential predictors: *speaker guise*, *speaker gender*, *speaker region, participant id*, *participant age of arrival to the U.S., participant gender,* and *participant social network strength*. The outcome variables were set to the ratings of the seven participant's impressions of the speakers: *formal, educated, warm, extroverted, intelligent, talks very fast,* and *engaging*. We used linear regression for modeling.

5. RESULTS

Between the seven impressions of social traits, only *formal* showed notable differences in the distribution of mean scores across all factor groups. The difference in the mean scores for *formal* was the greatest when interacting with the speakers' gender (see Table 2). The high range in the mean scores suggests that, overall, participants were least reluctant to express their opinions on the formality for speakers, while for the other social traits, they expressed mostly neutral opinions. Interestingly, while the scores for *educated* interacted with speaker *gender*, speaker *region*, (ING) *guise*, and self-reported *social network strength*, *intelligence* – a trait very close in meaning – did not show the same effects. Although the interaction of *educated* and *network strength* is statistically significant (see Table 2), the coefficient is only –.11. The network strength makes an impact on the directionality of the perception, but the effect is small. For both *educated* and *formal* (see Figure 1)*,* the 2 female speakers were rated higher than male speakers, regardless of (ING). The Southern speaker received higher average ratings for both *educated* and *formal* than the west coast speaker, revealing a reversed pattern from impressions previously documented for native English speakers [5].

Qualitatively, when performing the pilot study, we found the saliency of the distinction between velar and alveolar (ING) to vary by the participant's native dialect in Chinese, and the language proficiency in L2. Those who expressed having been occupied by the goal to comprehend the content did not notice the difference. Speakers that did not distinguish between alveolar and velar forms in their dialects found it more difficult to do so in English. Quantitatively in the experiment, the distinction between the (ING) variants did shape the pattern in the perception of *formal* and *educated*.

For both traits, the velar form received slightly higher mean scores than the alveolar form. While this pattern seems to match previous results [4,5], it should be noted that both forms received a quite neutral score (ranging from 2.21 to 2.94 on a 6.0 Likert scale), which suggests that (ING) is not highly indicative of the social meanings for the average Chinese listener. Finally, regardless of (ING) guises, participants with higher *social network strengths* were likely to give lower *educated* ratings to all speakers: perhaps, the more integrated individuals felt more confident in the host society and did not need to inflate their scores.

FORMAL ~ Speaker gender ($p < 0.0001$)		
Factors	Coefficient	Mean Score
female	0.747	3.08
male	–0.747	1.58
FORMAL ~ Speaker region ($p < 0.0001$)		
South	0.295	2.63
West Coast	–0.295	2.04
FORMAL ~ Speaker guise ($p = 0.04$)		
velar	0.125	2.46
alveolar	–0.125	2.21

EDUCATED ~ Speaker gender ($p < 0.0001$)		
Factors	Coefficient	Mean Score
female	0.506	3.36
male	–0.506	2.35
EDUCATED ~ Speaker region ($p < 0.0001$)		
South	0.286	3.14
West Coast	–0.286	2.57
EDUCATED ~ Social network score ($p = 0.002$)		
+1	–0.11	
EDUCATED ~ Speaker guise ($p = 0.02$)		
velar	0.0843	2.94
alveolar	–0.0843	2.55

Table 2. Variables significantly predicting the perception of *formal* (top) and *educated* (bottom).

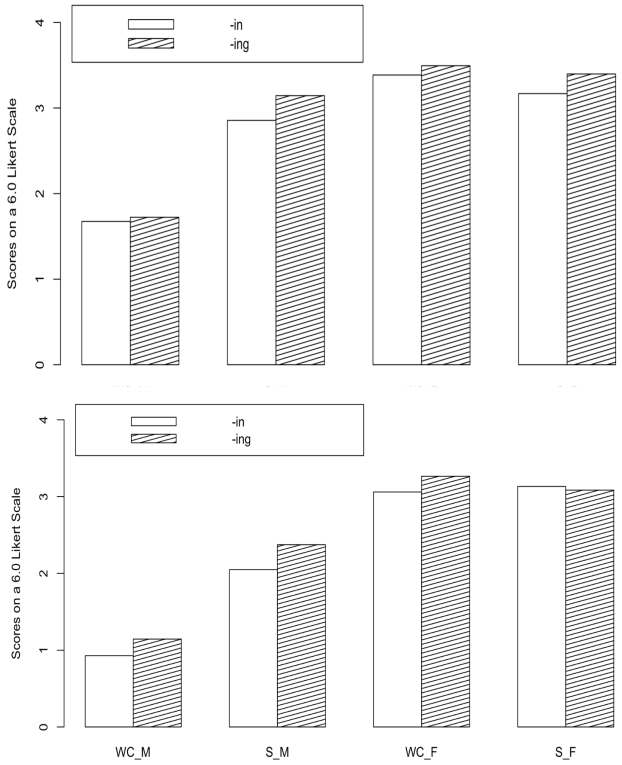

Figure 1. Average ratings for *educated* (top) and *formal* (bottom) across guises (*-in, -ing*) and speakers by gender (M, F) and region (West Coast: WC, Southern: S).

6. CONCLUSION

This work is the first attempt at investigating L2 listeners' social perceptions of (ING), a stable and socially meaningful English variable for L1 speakers. The study encompassed both a theoretical and a practical aim. For theoretical considerations, it aimed to provide a perspective on people's ability in sociopragmatic acquisition in L2, elucidating whether adults can acquire social meanings and nuanced linguistic variation at the phonological level in their second language. By co-investigating both L2 speakers' sociophonetic perceptions and their levels of interaction with the native speakers of the host community, we may eventually uncover the social conditions that facilitate the acquisition of sociopragmatic competence in their L2. At the practical level, when we establish that speech features at the phonological level can affect the way L2 speakers' perceive the social traits of the speakers, we can facilitate communication and enhance understanding in intercultural scenarios.

7. REFERENCES

[1] Labov, W. (1966). The Social Stratification of English in New York City. Washington, DC: Center for Applied Linguistics.

[2] Labov, W. (2001). Principles of linguistic change Volume 2: Social factors. *LANGUAGE IN SOCIETY-OXFORD-*, *29*.

[3] Trudgill, P., & Trudgill, S. (1974). *The social differentiation of English in Norwich* (Vol. 13). CUP Archive.

[4] Kiesling, S. F. (1998). Men's identities and sociolinguistic variation: The case of fraternity men. *Journal of Sociolinguistics*, *2*(1), 69-99.

[5] 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

[6] Schleef, E., Meyerhoff, M., & Clark, L. (2011). Teenagers' acquisition of variation. English World-wide, 32(2), 206-236. https://doi.Org/10.1075/eww.32.2.04sch

[7] Drummond, r. (2012). Aspects of identity in a second language: ING variation in the speech of Polish Migrants living in Manchester, UK. Language variation and change, 24(1),107-133.

[8] Milroy, L., & Margrain, S. (1980). Vernacular language loyalty and social network. *Language in Society*, 43-70.

[9] Milroy, L. (2000). Social network analysis and language change: introduction. European journal Of English studies , 4 (3), 217-223

[10] Ager, A., & Strang, A. (2008). Understanding integration: A conceptual framework. Journal of Refugee Studies, 21(2), 166-191.

[11] Plauen, E. O., & Song, Rui. (2019). *Fu yu zi quan ji: Cai se shuang yu ban = Father and son*. Beijing: Bei jing jiao yu chu ban she

[12] Boersma, P. & Weenink, D. (2009). Praat: doing phonetics by computer (Version 5.1.13)

[13] Lin, K. (2021, March). *Survey A*. Qualtrics. https://georgetown.az1.qualtrics.com/jfe/form/SV_exuqaUqfPk52lfM

[14] Lin, K. (2021, March). *Survey B*. Qualtrics. https://georgetown.az1.qualtrics.com/jfe/form/SV_PYfkRyRbwZXijc

[15] Qualtrics LLC. (2005). Qualtrics XM. https://www.qualtrics.com

[16] Johnson, D. E. (2009). Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis. *Language and linguistics compass*, *3*(1), 359-383.