THE ROLE OF IMMEDIATELY PRIOR EXPOSURE AND TALKER ACCENT ON SOCIOLINGUISTIC VARIANT IDENTIFICATION

Aini Li and Meredith Tamminga

Department of Linguistics, University of Pennsylvania
liaini@sas.upenn.edu; tamminga@ling.upenn.edu

ABSTRACT

In everyday speech perception, recognition of variable pronunciations is crucial, as different pronunciation variants may signal social information. Research in speech perception has investigated how perception can be robust to gradient phonetic variability, but less is known about how listeners integrate contextual cues in identifying discrete sociolinguistic variants. We present three experiments investigating the role of immediately prior exposure and its interaction with talker accent in perception of the English sociolinguistic alternation between *-ing* (/ɪŋ/) and *-in'* (/ɪn/). We hypothesize that having recently encoded one variant will increase the probability of perceiving the same variant subsequently during perception, and that this priming effect will be modulated by talker accent via social expectations. We find that the *-ing/-in'* variants can be primed in perception. However, this variant priming effect does not significantly interact with talker accent, suggesting that perception of variable pronunciations may not integrate local statistics and social unexpectedness simultaneously.

Keywords: speech perception, sociolinguistic variants, priming, talker accent

1. INTRODUCTION

Human speech is highly variable such that words can have different pronunciations. For instance, the word WALKING can be pronounced as either *walking* or *walkin'*. Crucially, these variable pronunciations often carry rich social information about the talker (i.e., age, gender, dialect, accent) [1, 2, 3]. A key challenge for listeners in everyday language use thus is to recognize different pronunciation variants reliably. Previous research in speech perception has examined why and how human speech perception, and by extension, sociolinguistic perception can be robust regardless of talker and speech variability [4, 5]. Recently, it has been suggested that speech perception can be viewed and modelled as a dynamic process of inference under uncertainty [6, 7, 8, 4]. That is, as speech signals unfold over time, listeners constantly draw linguistic inferences, update expectations, and generalize knowledge across words, sentences and talkers given their pre-existing knowledge of probabilistic distributions of acoustic, linguistic and social cues [5]. For instance, during processes such as phoneme categorization, when given an ambiguous sound halfway in the /s/-/ʃ/ continuum, listeners would be more likely to perceive the ambiguous sound as either /s/ or /ʃ/ based on what cue distributions they were exposed to and how they upweighted or downweighted certain cues. This approach offers a useful window to understand how listeners overcome gradient variability through cue integration during abstraction and generalization. However, we still have less idea of how listeners integrate different contextual cues in identifying equally well-formed discrete sociolinguistic variants, where the variants are distinct phonemes but are not lexically contrastive in the relevant context.

To probe this question, this study investigates how listeners recognize the well-studied sociolinguistic alternants *-ing* (/ɪŋ/) and *-in'* (/ɪn/) in English. We focus on the role of two different types of cues: immediately prior exposure and its interaction with talker accent. More specifically, we first test whether identification of discrete sociolinguistic variants can be primed by the immediately prior variant that the listener perceived (**Experiment 1**). In addition, we ask whether this priming effect, if it exists, can be modulated by talker accent via social expectation. Previously, it has been shown that linguistic expectations modulate priming effects in various experimental priming paradigms such as lexical repetition priming [9] and syntactic priming [10]. Of particular interest to our current investigation is the generalization that unexpectedness enhances priming; for example, unexpected verb constructions elicit stronger syntactic priming [10]. This raises the question of whether social unexpectedness might also enhance priming. We first ask whether a Southern accent in

American English influences how listeners identify *-ing*/*-in'* variants (**Experiment 2**). A Southern accent was chosen for this experiment because the variant *-in'* has been found to be strongly associated with the American South [11, 12, 13]. Despite intra-regional differences, Southern accents are among the most recognizable accents in the U.S. [14], which makes it easier for listeners to detect the difference between Southern and non-Southern accents. We then ask whether variant priming is modulated by talker accent (**Experiment 3**). Taken together, this study aims to use three experiments to develop a better understanding of how listeners integrate contextual cues in sociolinguistic variant identification.

2. THE CURRENT STUDY

2.1. Experiment 1: Priming variant choices in the lab

Exp. 1 tested the hypothesis that hearing one variant of *-ing*/*-in'* would make listeners more likely to perceive that same variant when given an ambiguous target for categorization. We compared whether a target word such as *sleepING*, with an ambiguous final nasal, is more likely to be categorized as *sleepin'* after *walkin'* than after *walking* (note that here we use <-ING> notation to indicate the ambiguous suffix form).

The experiment utilized a paradigm that combined two established tasks: lexical decision and forced-choice categorization. The lexical decision task was used to prime listeners' perception of different variants, while the forced-choice categorization task was used to probe listeners' perception of ambiguous targets. The ambiguity in the critical targets was created by performing source extraction on the *-in'* suffix through inverse filtering, resulting in the information filtered by the vocal tract, such as place of articulation, being masked while the information produced by the vocal folds, such as the intonational contour, remains unchanged. A norming study was used to estimate that the baseline *-ing* perception rate for these ambiguous items was 70%. The lexical decision task was adopted for prime trials to make it possible to filter out incorrect prime responses, which suggest the prime was not perceived as intended and therefore might not have been effective.

We adopted a within-subjects design with two critical conditions: *-in'*-primed and *-ing*-primed. To rule out the possibility that participants were converging to the model talker's overall *-ing*/*-in'* rate rather than being influenced by the intended prime, we controlled for the overall rates of the two variants throughout the experiment: whenever a participant responded to a target, they had been exposed to exactly 50% clear *-in* and 50% clear *-ing* primes from the model talker up to that point, regardless of which prime condition they were in.

Exp. 1 was implemented online in PCIbex [15]. 102 participants (self-reported monolingual American English speakers) were recruited from Prolific and received $10 for their participation (age range: 17-50y). Participants were informed that they would hear real English words and non-words (possible but non-existing), and that they needed to respond to these items in two different ways. On trials where they saw the prompt question, "Is this a word?", they needed to decide whether the word that they heard was a real English word or not by pressing a key (J for real words and F for non-words). On trials where they saw the prompt question, "Which word did you hear?", they needed to identify the word exactly the way they heard it by clicking on one of two options presented orthographically. To improve the acceptability of informal pronunciations in an experimental setting, following [16], participants were told that some of the words they would hear might be pronounced in a casual way, but that casually pronounced words were still real words of spoken American English. Participants were also given audio examples, then given a practice session with feedback (including words, non-words, and *-ing*/*-in'* variants) before the experiment.

38 critical ambiguous targets paired with 38 clear primes were created and included. Primes and targets were matched in whole word frequency using the SUBTLEX_{US} Log10CD measure [17]. Primes and targets also differed in the consonants immediately preceding the suffix *-ing*/*-in'*. 200 filler trials of various types including distractor sequences (e.g., sequences where targets after *-ing* or *-in'* did not have *-ing*/*-in'*. The word-to-nonword ratio for the lexical decision trials was 50%. All the experimental items were recorded by an adult white male native speaker of American English from New Jersey.

Analyses were conducted using R version 4.0.5 [18]. Mixed effects logistic regression was run using the *lme4* package version 1.1-27.1 [19], and plots were created using the *sjPlot* package version 2.8.9 [20]. Thirty-one participants were removed due to their accuracy in the lexical decision task being lower than 80%. Data from the remaining 71 participants were analyzed. Critical sequences where participants judged *-in'*-containing words as "nonword" were also excluded

from the final analysis. A mixed-effects logistic regression was conducted to predict listeners' *-ing*/*-in'* responses when identifying ambiguous targets (0 for *-in'*, 1 for *-ing*). Fixed effects were PRIME CONDITION (*-in'*-primed vs. *-ing*-primed), TARGET FREQUENCY, TRIAL NUMBER and the two-way interaction: PRIME CONDITION * TARGET FREQUENCY. PRIME CONDITION by SPEAKER and TARGET FREQUENCY by SPEAKER were included as random slopes correlated with a by-speaker random intercept. TARGET was also included as a random intercept. All the categorical predictors were sum-coded (-1, 1) and continuous predictors were scaled and centered using z-scores. Model output revealed a significant main effect of PRIME CONDITION ($\beta = 0.77$, $p < 0.001$), as Figure 1 illustrates.

Participants were significantly more likely to categorize an ambiguous target as containing *-ing* when they had just heard an *-ing* variant on the previous trial. No other predictors were statistically significant. On top of this predicted effect of prime, in both conditions, participants were more likely to perceive ambiguous targets as containing *-in'*, i.e., compared to the average *-ing* response baseline of 70% in the norming study (the red dashed line). It seems that when inferring which of two variants they heard, in the *-ing*-primed condition, listeners were matching their *-ing* responses to the model talker's overall *-ing* rates (the blue solid line), perhaps due to a global expectation of *-ing* from the experimental context. This could suggest that the priming effect is mostly driven by the *-in'*-primed condition, consistent with an account where unexpected forms elicit stronger priming.

2.2. Experiment 2: Talker accent modulates variant choices

Having established that phonological variant choices in perception can be primed, Exp. 2 tested whether listeners also integrated social information, specifically talker accent attributes, during discrete phonological variant categorization. In particular, we compared whether listeners would be more likely to perceive an ambiguous ING-suffixed word as *-in'*-containing, as opposed to *-ing*-containing, when the talker had a noticeable Southern accent instead of a less regionally-marked American English accent, which we refer to for convenience as a "general" accent.

Exp. 2 was implemented online through PCIbex. A total number of 102 participants (self-reported monolingual American English speakers) were recruited from both Prolific (N=80) and the undergraduate subject pool (N=32) in return for either monetary compensation (2 dollars) or course credit. The participant make-up from both sources was held consistent for each condition. Participants were randomly assigned to one of two critical conditions: a Southern accent condition and a general accent condition. In the Southern accent condition, all the items participants heard were Southern-accented. In the general accent condition, all stimuli were presented in a general accent instead. Participants were *not* given explicit information about what accent they would hear. The experiment used a simple forced-choice categorization task. For each stimulus item, participants were asked to identify whether they heard the word as containing *-ing* or *-in'*. The whole experiment lasted around 10 minutes.

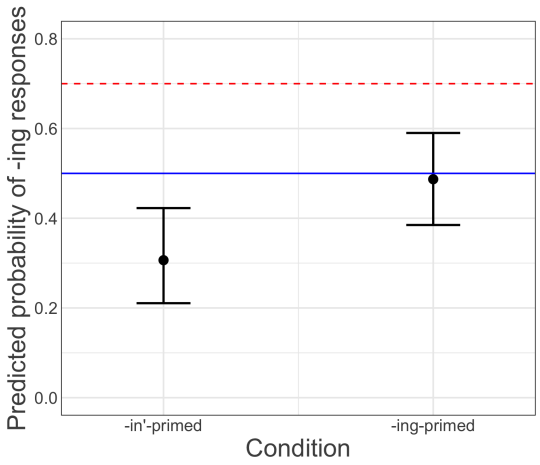

Figure 1: Predicted probability of perceived *-ing* responses in **categorization**, when all other predictors held at their average values, in probability terms. The red dashed line indicates the baseline *-ing* perception rates. The blue solid line refers to the overall *-ing* rates listeners were exposed to.

Target items (N = 38) from Experiment 1 were used as critical stimuli for this experiment. To keep the two accent conditions maximally similar, an adult white female bidialectal speaker from Louisiana who speaks both Southern-accented American English and a less-marked general American English was asked to produce critical items in both accents (henceforth: Southern guises vs. general guises). Ambiguous stimuli were created using the same procedure as in Experiment 1, with all ambiguous items derived from originally *-in'*-suffixed forms. In the end, 38 Southern-accented and 38 general-accented ambiguous targets were created and served as the final critical stimuli.

A mixed-effects logistic model was fit to predict

listeners' perceived *-ing* responses to ambiguous targets with TALKER ACCENT (General accent vs. Southern accent, sum-coded) as the independent variable. PARTICIPANT and TARGET were included as random intercepts. There is a main effect of TALKER ACCENT: Overall, listeners were significantly more likely to identify ambiguous targets as containing the *-ing* variant in the general accent condition, as opposed to the *-in'* variant in the Southern accent condition ($\beta = 1.30$, $p < 0.001$).

2.3. Experiment 3: Do variant priming and talker accent interact?

Exps. 1 and 2 showed independent effects of prior exposure and talker accent on variant perception. In Exp. 3 we ask whether these influences interact. If the priming in Exp. 1 partly reflected the unexpectedness of *-in'*, then the priming effect might be weaker in the Southern accent condition, as the association between *-in'* and Southern speech could make *-in'* primes less surprising.

For Exp. 3, 284 self-reported monolingual American English speakers were recruited from Prolific and the subject pool. As in Exp. 1, the two prime conditions (*-in'*-primed vs. *-ing*-primed) were manipulated within participants. Different from Exp. 1, talker accent was included as a between-participants factor. In the general-accent condition, participants were presented with general-accented stimuli (i.e. general-accented fillers, clear primes, and ambiguous targets), whereas in the Southern-accent condition, participants heard Southern-accented items. The same items from Exp. 1 were used as items for this experiment (both critical and filler items). Stimuli were produced in either a general accent or a Southern accent by the same bidialectal speaker from Exp. 2.

After excluding participants whose overall accuracy in lexical decision trials was lower than 80%, results from 155 participants were analyzed (general accent N = 83, Southern accent N = 72). A mixed-effects logistic regression was fit to predicted listeners' perceived *-ing* responses with the following fixed effects: TALKER ACCENT and PRIME CONDITION (in a two-way interaction), PRIME CONDITION and TARGET FREQUENCY (in a two-way interaction), and TRIAL NUMBER. PRIME CONDITION by SPEAKER and TARGET FREQUENCY by SPEAKER were included as random slopes and TARGET was included as a random intercept. Categorical predictors were sum-coded and continuous predictors were scaled and centered. We found a significant main effect of PRIME CONDITION, indicating that listeners on average were less likely to perceive *-ing* in ambiguous targets in the *-in'*-primed condition (β = -0.60, $p < 0.001$). There was also a significant effect of TALKER ACCENT, which indicates that listeners were significantly more likely to perceive *-ing* when the talker was general-accented ($\beta = 1.45$, $p < 0.001$). The interaction between PRIME CONDITION and TALKER ACCENT was not significant, which fails to support our hypothesis that the priming effect would be significantly weaker when the talker accent was Southern (β =-0.24, $p = 0.16$). Predicted effects of PRIME CONDITION and TALKER ACCENT are shown in Figure 2.

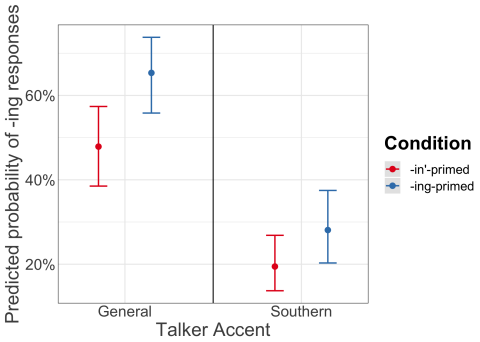

Figure 2: Predicted probability of perceived *-ing* responses across two talker accent conditions

3. DISCUSSION

We have demonstrated that first, the variant participants were recently exposed to can influence which variant they perceive subsequently. This is evidence that in speech perception, discrete sociolinguistic variants can be primed. The difference between the two conditions cannot be attributed to convergence towards the talker's overall *-ing*/*-in'* rate because the conditions do not differ in that rate. In addition, when listeners need to categorize ambiguous stimuli, talker accent influences their decisions as well. Even without being explicitly told that they would be hearing from a general or Southern-accented talker, listeners were able to infer this information from the acoustic cues in the stimuli, then use this information to associate variants with different talker accents based on participants' existing sociolinguistic knowledge (i.e., the relationship between regional accents and *-ing*/*-in'* variation). Finally, our results fail to provide evidence for a possible interaction between variant priming and talker accent, raising new questions about whether social unexpectedness modulates variant priming in the same way as linguistic unexpectedness does to other types of priming.

4. REFERENCES

[1] N. Niedzielski, "The effect of social information on the perception of sociolinguistic variables," *Journal of Language and Social Psychology*, vol. 18, no. 1, pp. 62–85, 1999.

[2] E. A. Strand, "Gender stereotype effects in speech processing," Ph.D. dissertation, Ohio State University, 2000.

[3] J. Hay and K. Drager, "Stuffed toys and speech perception," *Linguistics*, vol. 48, no. 4, pp. 865–892, 2010.

[4] D. F. Kleinschmidt and T. F. Jaeger, "Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel," *Psychological Review*, vol. 122, no. 2, pp. 148–203, 2015.

[5] D. F. Kleinschmidt, K. Weatherholtz, and T. Florian Jaeger, "Sociolinguistic perception as inference under uncertainty," *Topics in cognitive science*, vol. 10, no. 4, pp. 818–834, 2018.

[6] G. E. Peterson and H. L. Barney, "Control methods used in a study of the vowels," *Journal of the Acoustical Society of America*, vol. 24, no. 2, pp. 175–184, 1952.

[7] J. S. Allen, J. L. Miller, and D. DeSteno, "Individual talker differences in voice-onset-time," *Journal of the Acoustical Society of America*, vol. 113, no. 1, pp. 544–552, 2003.

[8] R. S. Newman, S. A. Clouse, and J. L. Burnham, "The perceptual consequences of within-talker variability in fricative production," *J Acoust Soc Am*, vol. 109, no. 3, pp. 1181–1196, 2001.

[9] D. L. Scarborough, C. Cortese, and H. S. Scarborough, "Frequency and repetition effects in lexical memory," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 3, no. 1, pp. 1–17, 1977.

[10] S. Bernolet and R. J. Hartsuiker, "Does verb bias modulate syntactic priming?" *Cognition*, vol. 114, no. 3, pp. 455–61, Mar 2010.

[11] A. Houston, "Continuity and change in English morphology: the variable (ING)," Ph.D. dissertation, University of Pennsylvania, 1985.

[12] W. Labov, *Principles of Linguistic Change. Vol. 2: Social Factors*. Oxford: Blackwell, 2001.

[13] K. Campbell-Kibler, "Accent, (ING), and the social logic of listener perceptions," *American Speech*, vol. 82, no. 1, pp. 32–64, 2007.

[14] R. Lippi-Green, "The standard language myth," in *English with an accent: language, ideology, and discrimination in the United States*. Routledge, 2012, ch. 4.

[15] J. Zehr and F. Schwarz, "PennController for Internet Based Experiments (IBEX)," Retrieved from https://doi.org/10.17605/OSF.IO/MD832, 2018.

[16] Y. White, "The processing and mental representation of ing variation," Ph.D. dissertation, University of Pennsylvania, 2021.

[17] M. Brysbaert and B. New, "Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English," *Behavioral Research Methods*, vol. 41, no. 4, pp. 977–90, Nov 2009.

[18] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2015.

[19] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[20] D. Lüdecke, P. D. Waggoner, and D. Makowski, "Insight: A unified interface to access information from model objects in r," *Journal of Open Source Software*, vol. 4, no. 38, p. 1412, 2019.