Factors Involved in Limited Convergence in (ING) Variation

Sarah Horwitz*

1 Introduction, Background and Hypotheses

Speakers have been observed to reuse recently articulated linguistic forms in cross-speaker conversational dynamics. One explanation for this behavior is convergence (moving closer to an interlocutor's speech), which is rooted in Communication Accommodation Theory and uses an identity-building framework to show how aspects of communication (phonetic, phonological, and morpho-syntactic elements of speech, as well as nonverbal cues) can index different intergroup identities (Giles and Ogay 2007, Gallois et al. 2005). Speakers can also diverge (move away from an interlocutor's speech) and practice maintenance (neither approach nor move away from an interlocutor's speech) (Giles and Ogay 2007, Gallois et al. 2005). In addition to socially-motivated convergence (*one* theory of convergence; the interactive alignment model and audience design are others), cognitive theories of repetition priming also explain cross-speaker behavior. Robust evidence supports the idea that speech perception and production are interdependent, a dynamic exemplified by recency effects in speech (Casserly and Pisoni 2010, Pickering and Garrod 2013 [in Pardo et al. 2018], Tamminga 2014, Pardo et al. 2018). Recency effects lead to priming, which in its most naïve sense means that hearing a form makes a speaker more likely to say that same form subsequently (Tamminga 2014, 2016). Superficially, priming and socially-motivated convergence effects present similarly, because both entail increasing similarity (and repetitiveness) in the variants speakers use during an interaction. Notably, only a socially-motivated theory predicts divergence.

The question at the core of our study asks if, during a sociolinguistic interview, an interviewer's use of the [ɪŋ] variant makes the speaker (interviewee) more likely to use [ɪŋ] in their immediately subsequent utterance, relative to their baseline likelihood of using [ɪŋ] (i.e. with no immediately preceding instance of ING from the interviewer). We do not expect to be able to tease apart specific socially-motivated convergence or priming influences in our study, and anticipate that patterns of results explainable by both accounts might support recent arguments that social dynamics and speech perception/production are connected, and so accounting for social factors is necessary in order to fully understand speech perception and production (i.e. Pardo et al. 2018, Campbell-Kibler 2010, Walker and Campbell-Kibler 2015, Abramowicz 2007). As previous work (Tamminga 2014) has investigated intraspeaker repetition in (ING), our study also provides a way to compare cross-speaker convergence effects with such evidence of intraspeaker persistence.

Variable (ING) (e.g. *workin' ~ working*; Tamminga 2014) is defined as the phonological alternation between the velar [ɪŋ] and apical [ɪn] variants. As variable (ING) is stable, and internally and externally conditioned, it provides a unique lens for simultaneously probing social and cognitive processes. (ING)'s "standard" [ɪŋ] variant is shown to be favored by speakers of higher social classes and by nominal (i.e. monomorphemic and gerund) tokens of (ING) (see Houston 1985, Abramowicz 2007, Campbell-Kibler 2007, Tagliamonte 2004, Tamminga 2014). The same sources argue that (ING)'s "non-standard" [ɪn] variant is favored by speakers of lower social classes and by verbal (i.e. progressive and participle) tokens of (ING). We predict that speakers will produce more [ɪŋ] following interviewer [ɪŋ], and more [ɪn] following interviewer [ɪn], because this would be consistent with both convergence and priming accounts. Our results will discuss this prediction in the context of cross-speaker vs. intraspeaker, gender, and grammatical effects.

2 Data and Methods

The natural language data of our study come from a roughly age- and gender-balanced 122-speaker subset of the Philadelphia Neighborhood Corpus (with 51 speakers born pre-1930, 46 between 1930-1959, and 25 post-1959; 53 male and 69 female speakers) (Labov and Rosenfelder

*I would like to thank the following people for their invaluable guidance and support: Dr. Meredith Tamminga and Dr. Lacey Wade; Dr. Julie Anne Legate and Ryan Budnick; my peers in Linguistics 300; Brandy Pineda; Angelica Revilla, Amanda Francis, and Tinyan Omere. Any remaining errors are my own.

2011). All speakers are white, and interviews lasted approximately 30-60 minutes and were conducted with a range of interviewers. Following Tamminga (2014:25,44), the interviewees' *and* interviewers' ING data were coded as "0" if they were pronounced with the apical [ɪn] variant, and they were coded as "1" if they were pronounced with the velar [ɪŋ] variant. Interviewer tokens occurring at most one minute before the subsequent speaker (ING) token were hand-coded in Praat (following Tamminga's 2019 findings suggesting priming effects are strong for about 30-60 seconds after a prime is uttered). All interviewee tokens were also coded into one of three categories: [ɪn]-primed, [ɪŋ]-primed, and unprimed. All tokens were coded for grammatical category (monomorphemic; gerund; participle; progressive; "-thing (ING)," which includes "something", "nothing", "anything," "everything" and was excluded from analyses due to the possibility of reduced articulation). Given limited token numbers, we only include monomorphemic (N=237; i.e. "pudding") and progressive (N=4407; i.e. "he is *singing*") tokens in our grammatical category analyses.

3 Results

3.1 Cross-speaker Convergence/Priming Effects vs. Intraspeaker Priming Effects

We first look at cross-speaker convergence/priming effects across all speakers, and find that the aggregate speaker mean [ɪŋ] rate for [ɪn]-primed tokens is **0.379** (N=116); the aggregate speaker mean [ɪŋ] rate for [ɪŋ]-primed tokens is **0.437** (N=458); and the aggregate speaker mean [ɪŋ] rate for unprimed tokens (which serves as an overall average [ɪŋ] baseline) is **0.444** (N=6966). That is, across all speakers, the [ɪn]-primed rate is slightly lower than the baseline (unprimed) [ɪŋ] rate, but the [ɪŋ]-primed rate is nearly identical to the baseline [ɪŋ] rate. This does not fully align with our expectations, and neither a convergence nor a priming theory can fully explain the pattern. Both convergence and priming accounts would predict that the [ɪn]-primed rate would be lower than the baseline [ɪŋ] rate, and that the [ɪŋ]-primed rate would be higher than the baseline [ɪŋ] rate. A divergence analysis can explain why the [ɪŋ]-primed rate is slightly lower than the baseline [ɪŋ] rate if we understand speakers to be taking an oppositional identity stance to the interviewer. However, given previous findings (Bock 1986, Pickering and Branigan 1999, Squires 2013 [all in Tamminga 2014]; Tamminga 2014) that attest strong priming effects, it is unpalatable to rely solely on social explanations. Further, the difference between the [ɪŋ]-primed rate and the baseline mean [ɪŋ] rate is so small that it is impossible to say if a true difference actually exists. Chi-square tests do not reveal significant differences between the aggregate [ɪn]-primed rate and the baseline [ɪŋ] rate, nor between the [ɪŋ]-primed rate and the baseline [ɪŋ] rate. We are thus unable to say what is causing these unexpected patterns, as well as uncertain if there are even any real effects at all.

Figure 1a presents the same cross-speaker convergence/priming effects at the individual speaker level (left):

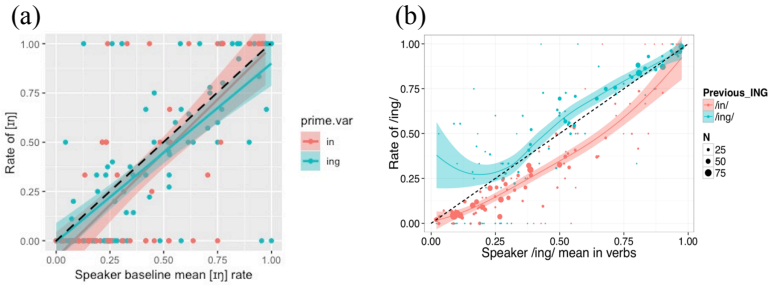

Figure 1a and b: Individual speaker values for (ING), after (a) [ɪn] and [ɪŋ] cross-speaker primes and (b) [ɪn] and [ɪŋ] intraspeaker primes (from Tamminga 2014:103).

In Figure 1a, we see that all three lines ([ɪn]-primed rate, [ɪŋ]-primed rate, and x=y) are essentially on top of each other, which indicates that there is hardly any effect of interviewer [ɪŋ] and [ɪn] primes on speakers' mean [ɪŋ] rates, *and* means that there is not a real difference between speaker [ɪn]-primed and [ɪŋ]-primed mean [ɪŋ] rates, and speaker unprimed mean [ɪŋ] rates.

While upon closer inspection, certain more nuanced patterns seem to emerge (i.e. for speakers with a relatively low baseline [ɪŋ] rate, [ɪŋ] primes seem to prime [ɪŋ] targets, and [ɪn] primes seem to prime [ɪn] targets; for speakers with a relatively high baseline [ɪŋ] rate, the pattern is reversed; and both the mean [ɪn]-primed [ɪŋ] rate *and* the mean [ɪŋ]-primed [ɪŋ] rate fall under our control dashed black line), it is uncertain if these patterns reflect any real differences, because the effects of the graph are small and insignificant. It is possible that more data would allow us to reliably detect the presence of small effects. Though we cannot draw firm conclusions about what is going on, we can compare these cross-speaker results with the strong *intraspeaker* persistence effects that Tamminga (2014) found in the same dataset, shown in Figure 1b (above, right).

Tamminga (2014) uses a speech persistence analysis similar to priming to explain the pattern visible in Figure 1b, suggesting that previous [ɪn] tokens prime subsequent [ɪn] tokens, and previous [ɪŋ] tokens prime subsequent [ɪŋ] tokens. Overall, it is surprising that there was not a stronger priming effect in our cross-speaker analysis, as there exist pervasive accounts of priming in the literature beyond Tamminga (2014) (i.e. Bock 1986, Pickering and Branigan 1999, Squires 2013 [all in Tamminga 2014]). At this time, we do not have enough knowledge to state whether priming (simple *or* complex) works differently within and across speakers. An alternate explanation of the behavior shown in Tamminga (2014) relies on theories of speaker stylistic sensitivity instead of priming. That is, the clustering of [ɪn] and [ɪŋ] tokens near each other would more reflect these adjacent tokens occurring in the same stylistic register (i.e. "casual" versus "careful", summarized in Labov 2001b) than it would reflect the speaker repeating the variant they most recently said. Yet while this purely social explanation may be valid, it is unappealing to explain linguistic behavior with only social factors, because so much other, robust evidence suggests that priming phenomena *do* occur in speech (Tamminga 2019). It appears our cross-speaker convergence and Tamminga's (2014) intraspeaker persistence are different; but we cannot make any further claims now.

3.2 Effect of Gender and (ING) Grammatical Category on Convergence/Priming Patterns

Figure 2a (below, left) presents aggregate speaker mean [ɪŋ] rates for [ɪn]-primed, [ɪŋ]-primed, and unprimed (ING) tokens, with speakers separated by gender:

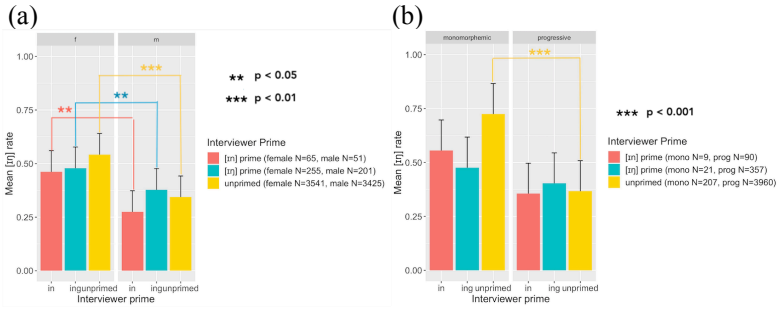

Figure 2a and b: Aggregate speaker mean [ɪŋ] rates for [ɪn]-primed, [ɪŋ]-primed, and unprimed (ING) tokens, by (a) gender and (b) token grammatical class.

We see that among male speakers, the [ɪn]-primed rate is less than the baseline [ɪŋ] rate, which is less than the [ɪŋ]-primed rate; whereas among female speakers, the [ɪn]-primed rate is less than the [ɪŋ]-primed rate, which is less than the baseline [ɪŋ] rate. This general pattern among males is as expected, following both priming and convergence analyses. The pattern among females does not align with a priming analysis, as their [ɪŋ]-primed rate is lower than their baseline [ɪŋ] rate, though it is possible that females are diverging from the interviewer's speech. Chi-square testing reveals statistically significant differences in the mean [ɪŋ] rates between female and male speakers. These differences tell us that males use less of the prestige [ɪŋ] variant, which aligns with well-established variationist patterns summarized by Labov's (2001a) "Principle I" of language change (the gender paradox). However, this gender effect is not enough to support calls for a more integrated model of social cognition (i.e. Campbell-Kibler 2010, Pardo et al. 2018).

Figure 2b (above, right) presents the [ɪn]-primed, [ɪŋ]-primed, and baseline mean [ɪŋ] rates for monomorphemic and progressive (ING). A chi-square test reveals that the monomorphemic unprimed mean [ɪŋ] rate is significantly higher than the progressive unprimed mean [ɪŋ] rate ($p < 0.001$), which is expected given that monomorphemic (ING) has been shown to favor the [ɪŋ] variant. The lack of other significant differences is not shocking due to the small number of monomorphemic [ɪn]-primed and [ɪŋ]-primed tokens (N=9 and N=21, respectively). Given this lack of significance and the small sample size, we cannot be confident that this visual pattern reflects what is truly happening in the data. Turning our focus to progressive (ING), we do not find significant differences between the [ɪn]-primed, [ɪŋ]-primed, and unprimed mean [ɪŋ] rates, and thus cannot draw conclusions about grammatically conditioned convergence/priming behavior. An additional analysis of the effect of progressive [ɪŋ] and [ɪn] primes on progressive and monomorphemic (ING) targets did not reveal statistically significant differences between the [ɪŋ]-primed and [ɪn]-primed mean [ɪŋ] rates of progressive prime/target pairs, *nor* between the [ɪŋ]-primed and [ɪn]-primed mean [ɪŋ] rates of progressive prime/monomorphemic target pairs. This is not shocking given our small number of tokens, meaning it is possible that with more tokens, statistically significant differences would emerge, and with them additional support for recent arguments that (ING) may not actually be one monolithic variable (i.e. Tagliamonte 2004, Tamminga 2016).

3.3 Regression Analysis

To substantiate the above analyses, we ran a linear mixed-effects regression to evaluate our cross-speaker convergence trends in a more statistically powerful way that allowed multiple predictors to be included simultaneously. To account for different speakers having different baseline [ɪŋ] rates as well as a different number of tokens, our model included a random by-speaker intercept. The results of our regression (shown below in Table 1) support the findings presented above: we do not find a significant difference between [ɪŋ]-primed [ɪŋ] rate and overall mean [ɪŋ] rate, and though we find that women's [ɪŋ] rates are significantly higher than the mean of men's and women's [ɪŋ] rates, we cannot say if the general convergence and gender differences are independent of each other, because our (ING) prime and gender interaction proved insignificant.

	Estimate	Std. Error	z value	p-value
(Intercept)	-0.719599	0.244477	-2.943	0.00325 **
Prime variant [ɪn]	0.005537	0.153206	0.036	0.97117
Gender (female)	0.510217	0.242831	2.101	0.03563 *
Prime variant [ɪn] * gender (female)	0.081035	0.152473	0.531	0.59509

Table 1: Linear Regression Output (sum-coded, log-odds); * = $p < 0.05$ and ** = $p < 0.01$.

4 Results and Conclusion

We were surprised to find null effects of interviewer [ɪŋ] and [ɪn] primes on speaker mean [ɪŋ] rates, relative to speakers' baseline mean [ɪŋ] rates. While a priming approach does not give us a way to explain why interviewer [ɪŋ] and [ɪn] tokens did not seem to prime speaker [ɪŋ] and [ɪn] tokens, socially-motivated convergence does offer some theoretical insights into these patterns. Following a socially-motivated convergence perspective, it is possible that speakers are generally diverging from instances of interviewer [ɪŋ] to produce more [ɪn] tokens; thus, some sort of social information could be interfering with a straightforward recency (priming) effect. Yet because the effects we see are so small, we cannot be confident that they reflect what is really happening in the data, nor can we be confident they would be substantiated with additional data. By comparing our lack of cross-speaker convergence effects with previous evidence of robust intraspeaker persistence effects (Tamminga 2014), it appears that cross-speaker convergence and intraspeaker persistence are different. Subsequent analyses of the effect of speaker gender on convergence/priming patterns, as well as of the effect of progressive [ɪŋ] and [ɪn] primes on progressive and monomorphemic (ING) target tokens' mean [ɪŋ] rates, seem tentatively promising but ultimately proved insignificant. Furthermore, while we see that monomorphemic (ING) tokens appear to have a higher mean [ɪŋ] rate than progressive (ING) tokens, a lack of other significant results and a generally small amount of data prevent us from drawing any conclusions on this front. While we are unable to make any firm claims

based on the overwhelmingly null results of our study, the notion of social and cognitive factors interdependently affecting speech perception and production remains interesting. Ultimately, our results underscore the need for additional corpus studies to approach topics of convergence and repetition priming. They also suggest that it could be interesting for further work to look more deeply into speech repetition effects within and across speakers. As with any study, extensions of this one would benefit from additional data, which would allow us to investigate by-speaker baseline mean [ɪŋ] rates with more confidence and allow more comprehensive grammatical category analyses to be conducted.

References

Abramowicz, Lukasz. 2007. Sociolinguistics meets Exemplar Theory: Frequency and recency effects in (ing). In *U. Penn Working Papers in Linguistics 13.2: Selected Papers from NWAV 35*, ed. S. Brody, M. Friesner, L. Mackenzie, and J. Tauberer, 27–37.

Aguilar, Lauren., Geraldine Downey, Robert Krauss, Jennifer Pardo, Sean Lane, and Niall Bolger. 2016. A Dyadic Perspective on Speech Accommodation and Social Connection: Both Partners' Rejection Sensitivity Matters. *Journal of Personality* 84.2:165-177.

Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American Speech* 82.1:32-64.

Campbell-Kibler, Kathryn. 2010. New Directions in Sociolinguistic Cognition. *U. Penn Working Papers in Linguistics 15.2: Selected Papers from NWAV 37*, ed. K. Gorman and L. MacKenzie, 31-39.

Casserly, Elizabeth D. and David B. Pisoni. 2010. Speech perception and production. *Wiley Interdisciplinary Reviews Cognitive Science 1*.5:629–647

Gallois, Cindy, Tania Ogay, and Howard Giles. 2005. Communication Accommodation Theory: A look back and a look ahead. In *Theorizing about intercultural communication*, ed. W.B. Gudykunst, 121–148. Thousand Oaks: Sage.

Giles, Howard, and Tania Ogay, T. 2007. Communication Accommodation Theory. In *Explaining communication: Contemporary theories and exemplars*, ed. B. B. Whaley and W. Samter, 293–310. Mahwah, NJ: Lawrence Erlbaum.

Houston, Anne. 1985. Continuity and Change in English Morphology: The Variable (ING). Doctoral disserta tion, University of Pennsylvania.

Labov, William. 2001a. Principles of Linguistic Change: Volume 2, Social Factors. Blackwell Publishers.

Labov, William. 2001b. The anatomy of style-shifting. In *Style and Sociolinguistic Variation*, ed. P. Eckert and J.R. Rickford, 85-108. Cambridge: Cambridge University Press.

Labov, William and Ingrid Rosenfelder. 2011. The Philadelphia Neighborhood Corpus of LING 560 studies, 1972-2010. With support of NSF contract 921643.

Pardo, Jennifer. 2006. On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119:2382–2393.

Pardo, Jennifer. 2013. Measuring phonetic convergence in speech production. *Frontiers in Psychology* 4:1–5.

Pardo, Jennifer, Adelya Urmanche, Sherilyn Wilman, Jaclyn Wiener, Nicholas Mason, Keagan Francis, and Melanie Ward. 2018. A comparison of phonetic convergence in conversational interaction and speech shadowing. *Journal of Phonetics* 69:1-11.

Tagliamonte, Sali A. 2004. Someth[in]'s go[ing] on!: Variable ing at ground zero. In *Language Variation in Europe: Papers from the Second International Conference on Language Variation in Europe*, ed. B. L. Gunnarsson et al. Department of Scandinavian Languages, Uppsala University.

Tamminga, M. 2014. Persistence in the Production of Linguistic Variation. Doctoral dissertation, University of Pennsylvania.

Tamminga, M. 2016. Persistence in phonological and morphological variation. *Language Variation and Change* 28:335-356.

Tamminga, M. 2019. Sources of Microtemporal Clustering in Sociolinguistic Sequences. *Frontiers in Artificial Intelligence* 2:10.

Walker, Abby and Kathryn Campbell-Kibler. 2015. Repeat what after whom? Exploring variable selectivity in a cross-dialectal shadowing task. *Frontiers in Psychology* 6:546.

Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104–6305
shorwi@sas.upenn.edu