Sociolinguistic Variation and Language Processing (S-VALP)
The Inn at Virginia Tech, Blacksburg, VA.
March 31st – April 2nd, 2016.

Conference Abstracts

Thanks to the conference sponsors:
NSF SBE (Linguistics) #1530780
Virginia Tech Department of English
Virginia Tech Department of Foreign Languages
The Inn at Virginia Tech
Do social expectations towards speakers affect perception of single words?

Previous studies of the effect of face and voice congruency on comprehension have either concluded that a loss of comprehension stems from prejudice against certain ethnicities (Rubin 1992) or that it is merely the result of incongruency between the ethnicity and accent of the speaker (McGowan 2015). This study uses an auditory lexical decision task with visual priming to examine the effect of congruency between an auditory and a visual input on the processing speed of single words. We examine the effect of standard vs. non-standard stimuli when paired with either the face of a white person or a person of colour. The auditory stimuli are taken from two registers of Copenhagen Danish, “modern” (standard) vs. “street” (non-standard). These registers have been shown in production and language attitude studies to be associated with white and coloured speakers respectively (Maegaard 2007; Pharao et al. 2014). The stimuli in this study were constructed by manipulating young male speakers’ natural word initial productions of /s/ and /t/ into either [s] and [tʰ] (“street”-variants) or [s] and [tʰ] (“modern” variants). If the coloured face facilitates recognition of words with ”street” variants but not the words with “modern” variants and vice versa for the white face this will support a congruency explanation. If the white face facilitates perception regardless of the spoken variant, whereas the coloured face inhibits it, this will indicate that it is not a matter of congruency but rather an effect of prejudice detached from any detailed expectations about which speakers use the two registers.

References:
Hyperarticulation in Mothers’ Speech to Babies and Puppies

**Background and Aims:** Several studies have found that mothers’ hyperarticulate their infant-directed speech (IDS), primarily by exaggerating vowels. Importantly, such hyperarticulation may increase perceptual word clarity, and is positively correlated with infants’ vocabulary size. Moreover, mothers’ IDS is hyperarticulated in comparison to speech to adult pets (PDS), even though both IDS and PDS are higher pitched, and rated as higher in affect than adult-directed speech (ADS). Mothers may hyperarticulate IDS because they assume infants are language-learners, whereas pets are not. However, to fully evaluate this language-recipient hypothesis, maternal IDS should be compared to both human and non-human infants. We contrasted mothers’ IDS to their 6-month-olds, to puppies (PupDS), and to adults. If the language-learning status evokes hyperarticulation, then only IDS should be hyperarticulated.

**Method:** To date, six mothers have been recorded speaking to their infant, a young puppy, and an adult (counterbalanced for order) while referencing 3 objects: a small boot, a small box, and a bead (representing three English corner vowels). Clearly enunciated words (boot, box, bead) were then excised using Adobe Audition, and their vowels (10 msec into the start of and before the end of voicing) were isolated using Praat software. The first (F1) and second (F2) formants of 282 vowels (92 boot/95 box/95 bead) have been analyzed at start, center, and end positions.

**Key Findings:** Comparisons of F1 and F2 for each word have revealed that IDS and PupDS did not differ significantly at any location, for any vowel. In contrast, they differed significantly from ADS depending on location and formant: Both IDS(F1) and PupDS(F1) were different from ADS(F1) at the start of /u/ and center of /i/; they differed from ADS at all three locations in /i/, the end of /u/, and the center and end of /a/.

**Conclusion:** In these recordings, hyperarticulation did not differentiate between language-learning status of the infant targets (human or canine); both types were more hyperarticulated than ADS. This raises the possibility that maternal hyperarticulation may be more related to emotional expressiveness than to a desire to teach language. We are currently rating IDS, PupDS, and ADS sentences (filtered to remove content) for emotional valence and expect that IDS and PupDS will not differ, but will both higher in affect than ADS. We are also collecting speech samples from additional mothers to increase sample size, and will include some adult dogs to test this hypothesis further.
Phonetic Accommodation: Who is speaking and are we really listening?

Convergence in spoken language -- the act of one’s speech patterns becoming more similar to speech one has been recently exposed to -- has been approached from a range of perspectives over the course of several decades. Social psychologists have argued that convergence is socially moderated, with its goal to facilitate social interaction and decrease social distance (e.g., Giles, 1973). Psycholinguists and speech researchers have focused their inquiries into the phenomenon in terms of the cognitive organization of spoken language and how the process indicates a coupling between speech perception and production (e.g., Goldinger, 1998; Sancier & Fowler, 1997). Examining convergence with both of these research perspectives in mind (e.g., Pardo, 2006) suggests that convergence is dictated by a confluence of social and cognitive forces. Using this multidisciplinary approach, in this talk I focus on spontaneous phonetic accommodation -- a highly non-social lab-based auditory-naming paradigm that, nevertheless, illustrates that speech perception and production are inherently social activities. Listeners’ degree of accommodation appears to be partially governed by social preferences (Babel, 2010; 2012; Babel, McAuliffe, & Haber, 2014) and social evaluation of voices (Babel, McGuire, Walters, & Nichols, 2014), in addition to non-social factors related to a language’s phonological inventory (Nielsen, 2011) and lexical frequency (Goldinger, 1998). Voices inherently carry social signals and thus social meaning, and as listeners evaluate the messenger this evaluation affects how listeners understand, assess, and evaluate the message itself (e.g., Sumner et al., 2014). In discussing this work I also present supporting evidence from face recognition and visual processing (Bernstein et al., 2007; Todd et al., 2012), which suggests our visual perception systems can be selective and enhance salient experiences. Auditory and linguistic processing are likely affected by processes of selective attention and enhancement as well. Together, this work indicates our theories of spoken language need to be mindful of the many ways in which we listen, evaluate, and produce speech.
Proactive inhibitory control and the adoption of sound changes in-progress

In order to learn to speak a given variety and integrate into that speech community, one must unravel the probabilistic tendencies governing its often sizable variation. This is further complicated by incipient sound change, which alters these tendencies as a function of a number of internal and external factors. Sociolinguistic factors such as age, ethnicity, and socioeconomic status modulate one’s input by constraining one’s communicative network, but an individual is also limited in real time by his or her cognitive capacity and processing style (cf. Yu 2013). For example, proactive inhibition is utilized to plan for upcoming events and, by its nature, is subject-internal and based on one’s own expectations about future input (cf. Braver 2012; Braver et al. 2007). An individual who relies on proactive control may appeal to previously-encountered rather than novel categories when processing speech, and might consequently pay less attention to subtle variations in the acoustic signal of his/her interlocutor. For sound changes in progress, new variants gradually increase in frequency relative to traditional variants, requiring him/her to also adapt to new distributional patterns.

Changes in Philadelphia’s phonology over the last century have been studied in depth, at least with respect to its White population (cf. Labov et al. 2013). White speakers have developed a split /æ/ system, wherein /æ/ tenses and diphthongizes to [ea] in closed syllables with coda voiceless fricatives or nasals. Originally, this system mirrored that of New York, but it has since been restricted in younger and more educated speakers to only those closed syllables with nasal codas due to stigmatization and the feature’s association to lower class New York speech (Labov et al. 2013). Contrariwise, Canadian Raising (raising /aj/ to [aj] in syllables with voiceless codas) is a rapidly developing change-in-progress which shows little stylistic stratification (Labov et al. 2013).

This study predicts degree of separation between traditional and novel variants for Canadian Raising and the Split-æ system in bilingual Puerto Ricans living in Philadelphia (N=11) as a function of age and proactive inhibitory control (via the AX-CPT task; cf. Cohen et al. 1999; Braver et al. 2007), using data from sociolinguistic interviews that were force-aligned using FAVE (Rosenfelder et al. 2011; N_{split-æ}=2,385 and N_{CanRaise}=2,369). To measure a speaker’s distinction between the two variants of each phoneme, the Pillai score (Hay et al. 2006; Nycz and Hall-Lew 2013) is used as a dependent variable. Preliminary results indicate that a phoneme’s social stratification modulates adoption of the novel variant (see Figure 1): younger speakers show a strong distinction between the novel raised [aj] and traditional [aj], but demonstrate almost no distinction between [æ] and its marked, tensed variant [ea] (older speakers don’t tend to adopt novel variants to begin with). Additionally, increased proactive inhibitory control appears to make speakers more resistant to novel variants, independent of their social stratification (see Figure 2).

These results are discussed within the scope of other research being conducted on sound change and language processing, and suggestions for conducting socially-informed psycholinguistic research are offered.
Figure 1:
Age Effects in Adoption of Novel Variant

Figure 2:
Proactive Control Suppresses Adoption of Novel Variant
References:


Using naturalistic and engaging tasks to measure phonetic convergence

The acoustic cues used to signal phonological contrasts vary across languages, dialects, and even between individual talkers. Voice onset time (VOT), for example, has well known patterns of cross-linguistic variation (Lisker & Abramson, 1964), but there is also evidence that it varies between men and women (Ryalls et al., 1997). Interlocutors must overcome this variability in order to successfully communicate with each other. One mechanism that may allow them to do so is phonetic convergence, a process by which conversational partners adjust their speech patterns to sound more like each other.

Despite the potential utility of this strategy, the extent to which interlocutors converge is affected by a number of factors (e.g., Pardo, 2006). Moreover, it may be difficult to elicit convergence using typical laboratory approaches (Olmstead et al., 2013). For example, imitation tasks requiring participants to produce non-native phones may indicate limits on convergence capability or ability to imitate non-native speech sounds given the talkers’ phonemic inventory. Alternatively, such tasks maybe too artificial to examine whether convergence occurs. This points to a more general challenge in studying speech communication: researchers need tasks that elicit naturalistic language use in laboratory settings (Toscano et al., 2015). Studying speech production in a more engaging, naturalistic task may provide information about the circumstances that lead interlocutors to converge, or in some cases, diverge.

We addressed this issue using a game-based approach that provides an engaging, yet experimentally controlled, task. Pairs of participants worked together to complete puzzles in the computer game Minecraft. Critical puzzles consisted of mazes, with the correct route being determined by minimal pairs (/b,d,g/ vs. /p,t,k/ distinctions) displayed on signs in the maze. For instance, if a participant needed to determine whether they should go left (as indicated by a sign with the word bear) or right (as indicated by the word pear), they would receive the appropriate direction from their partner. Thirty word-initial voicing minimal pairs were chosen to create conditions in which participants’ ability to understand each other depends on fine-grained phonetic cues that may differ between talkers. We examined the degree to which participants converged over the course of the one-hour experiment by measuring VOTs as a function of when they were produced in the game.

Similar to previous studies, we found that some subjects showed phonetic convergence, whereas others diverged. One factor that was related to resulting convergence is the amount of time it took to complete the experiment. Interestingly, subjects who completed it more quickly were less likely to converge (suggesting successful communication without a need to converge to each other’s speech patterns). Finally, we also examined the overall distribution of VOT values. Mean VOT values for voiced and voiceless stops were longer than what has been previously reported (voiced VOT: 25 ms; voiceless VOT: 101 ms; cf: Lisker & Abramson, 1964), suggesting that the overall distribution of acoustic cues is task dependent. Together, these results suggest that engaging, naturalistic tasks may yield results that more accurately reflect real-world phonetic variation than traditional laboratory experiments.

References


Primming and sociophonetic variation in natural speech

Understanding how and why pronunciations vary and change has been a dominant theme in variationist sociolinguistics since the 1960s (1966, 1994, 2001, 2010). Linguistic variability has also been a core area of focus in psychology and cognitive science. Work on structural priming, for example, has shown repeatedly that where variation exists in language, an alternative form, once used, persists in working memory and has a greater chance of reuse next time (Bock 1986; Bock & Loebell 1990; Branigan et al 2000). While there have been efforts to connect priming research with sociolinguistics at the level of grammar (Poplack 1980; Travis 2007), there has been almost no work considering the potential role of structural priming as a motivating factor in studies of accent variation and change (although see recently Clark (2014) and Tamminga (2014)).

This paper explores the role of structural priming at both the phonetic and phonological level in a socially constrained sound change in progress. The data come from the QuakeBox corpus, a collection of 723 earthquake monologues told by people recounting their memories of the Canterbury earthquakes of 2010-11 (Walsh et al 2013). Here I examine variation in the realization of over 5000 tokens of word medial, intervocalic /t/ in New Zealand English. This phonological environment provides the opportunity for a wide range of allophonic variation which is known to pattern by speaker sex - males prefer the voiced variants ([d] and [ɾ]) and females prefer the voiceless variants ([t], [ts], [s]; Hay and Foulkes, in press). By fitting a mixed effects logistic regression model to the data (Baayen et al 2008), incorporating both the realisation of the preceding variable, and a range of other linguistic and social predictors, I show two main findings:

1. Phonetic variants tend to cluster together in naturally occurring speech, and show the same tendencies as grammatical variables in studies of structural priming. That is there is a stronger tendency for similar variants to occur together in running speech when the preceding and following instance share the same lemma (similar to the ‘lexical boost’ effect reported in Hartsuiker et al., 2008 and Jaeger & Snider 2013), and this effect diminishes as the time between the instances increases. In this instance, synthesising research in cognitive science and psychology with sociolinguistics provides us with a better understanding of the factors underpinning a sound change in progress.

2. This priming effect also interacts with known and widely attested sociolinguistic predictors of variation in this community. Females show more priming with [s] and males show more priming with [ɾ]. Studies of syntactic priming have also found asymmetrical priming, with some variants more likely to show priming than others (Bock, 1986; Ferreira, 2003; Hartsuiker & Kolk, 1998; Kaschak et al 2011; Scheepers, 2003), but this is typically assumed to behave in the same way across the participant pool. By incorporating insights from sociolinguistics into cognitive science and understanding how this sound change is progressing through the speech community, we can see that priming may not behave in the same way across all speakers but is itself socially conditioned.
References
Social Meaning and Less Conscious Sociolinguistic Features: The Case of /u/

According to Labov (1965), a sociolinguistic variable becomes a stereotype when it becomes the “overt topic of social comment” (112). Although recent research in sociolinguistic production and meaning has looked at non-stereotype features, such as falsetto (Podesva 2007), many of the recent studies in sociolinguistic perception have yet to turn their attention to features that speakers cannot directly comment on (Campbell-Kibler 2010 and Walker et al. 2014, but see also Labov 1966). The current study seeks to fill this gap in the literature by asking: do less-conscious sociolinguistic features, including those without colloquial labels, carry meaning in perception and is that meaning modulated by perceived talker identity?

In order to answer these questions, a matched guise experiment (Lambert et al. 1960) was conducted, testing for social meaning with front and back /u/. /u/ was hypothesized to not be a stereotype under Labov’s (1965) definition. A total of 12 talkers were chosen, which were equally and orthogonally divided by race (Black and White) and gender (male and female). One token of /u/ from each talker’s read speech sample was then manipulated to back and front positions relative to the talker's vowel space. Ninety-four participants at a large Midwestern university then listened to these 24 recordings (12 talkers x 2 guises) and rated them on several six-point scales, including how feminine, intelligent, urban, and trendy each recording sounded. Ratings that loaded together in a factor analysis were combined and then modeled, with fixed effects for talker race, talker gender, /u/ guise, and their interactions; random intercepts for talker and listener; and random slopes for gender, race, and /u/ over listener.

Results show that a fronted /u/ sounded more feminine ($p=0.0015$), more White ($p<0.0001$), and more country ($p=0.0001$). Moreover, listeners evaluated fronted /u/ as more intelligent and educated than back /u/ ($p<0.0001$). However, there was also a significant race-gender-/u/ interaction ($p=0.0008$), such that Black male talkers were not rated as more educated and intelligent in the fronted /u/ condition. Additionally, listeners evaluated recordings with fronted /u/ as more trendy ($p=0.0031$), but only for White talkers ($p=0.0072$).

In the post-survey questionnaire, participants experienced difficulty discussing /u/, as none mentioned any colloquial labels for distinguishing between variants. Only one-third of participants reported paying attention to the word containing the /u/ vowel and participants had trouble ascribing front and back variants to different regions. One participant in particular said front /u/ is said in New England and Alaska, which are highly different dialect regions.

Taken with the responses from post-survey interviews, the data suggest that like meanings assigned to highly conscious features, meanings assigned to features lower in explicit awareness are not fixed or uniform in perception. In the case of /u/, listeners rated fronted /u/ as trendy for White talkers, but not for Black talkers. This study thus supports the idea that listeners can make use of features low in explicit awareness and/or that lack colloquial terms in order to form impressions.
References


Regional Biases and Phonetic Imitation

The process of phonetic imitation, in which speakers take on acoustic characteristics of their interlocutor’s speech, has been characterized by some as a completely automatic process that occurs as a result of phonological storage and processing (Goldinger, 1998). Others, however, have argued that the extent of imitation can be influenced by linguistic or social factors (Babel, 2012). This study examined the effect of social information on phonetic imitation by investigating how attitudes towards a speaker’s region of origin affected the degree to which that speaker’s vowels were imitated.

Forty-eight American English speakers from the Midland dialect region participated in a shadowing task in which they read words aloud and then repeated the same words produced by a model talker. The model talker’s speech contained features of Southern American English, which is subject to negative biases in other regions of the country (Preston, 1997). Participants were given different information about the regional background of the model talker. In two “Southern-labeled” conditions, participants were either told that the talker was from Eastern Kentucky or Savannah. While both of these locations are Southern, they differ in terms of the specific stereotypes associated with them (e.g. Baird, 2014). In a third condition, participants were told the talker was from Columbus, Ohio, a Midland city likely to be considered local in terms of regional identity. A fourth group of participants were given no information on the regional origin of the model talker. Following the shadowing task, participants completed a survey about their perceptions of the talker and the region they believed she was from.

Production data were collected on several vowels, including /ɔs/, /u/ and /æi/, which have characteristic Southern pronunciations (Labov, Ash, and Boberg, 2006). Read and shadowed productions were compared to the model talker’s production in order to assess whether or not imitation occurred. Analyses revealed significant imitation of the F2 of /ɔs/ and trajectory length of /æi/, and found that /ɔs/ was imitated more than the other vowels. I will argue that these between-vowel differences may be due in part to linguistic factors and in part to the relative level of social salience and stigma associated with the Southern variants of these vowels. The survey data revealed the expected stereotypes and relative attitudes towards the different regions, but an overall between-conditions effect on imitation was not found.

Based on these results, it appears that perceptions of the talker and her speech may have been influenced by general regional stereotypes, but that imitation of the individual vowel features may have been affected by phonetic and social factors related to the vowel variants themselves. As with previous research, the results of this study suggest that the imitation process is not completely automatic, and may be influenced by certain language-external social information. This, in turn, suggests that the cognitive mechanisms underlying linguistic processing must be linked to social information.
References


Moving variation up the food chain.

With its roots in dialectology and sound change, the study of variation has focused on a limited set of segmental variables, selected by virtue of their status as clear markers of speakers’ regional origins or membership in macro-social categories (e.g. class, gender). In recent years, a variety of ethnographic (e.g. Eckert 2008) and case (e.g. Podesva 2007) studies have broadened the view of systematic variability beyond the segmental to prosody and voice quality. Furthermore, a closer view of situational shifts (e.g. Eckert 2011; Podesva 2007; Moore and Podesva 2009) shows that linguistic features can express stance and affect, and that the meanings of variables are not fixed, but emerge through their participation in styles. The meanings of variables, thus, cannot be predicted from their macro-correlations; rather, the macro-social patterning of variation reflects the patterning of the expression of more personal meaning in a stratified society. This talk will argue that the potential to convey social information non-propositionally is a design feature of language, and that variation is the primary locus of this potential. Thus socially meaningful variation is not just a by-product of linguistic change and contact, but an integral part of the semantics of language.


Effects of language background on the development of sociolinguistic awareness: accent processing in monolingual and multilingual 5-6 year old children.

Complex international migration patterns have led to increasingly diverse multidialectal and multilingual communities, particularly within large urban centres such as London, UK (e.g., Vertovec, 2007). Such complex migration patterns mean that native, monolingual, children are likely to encounter not just different native regional accents but also foreign-accented speech. Children raised bi- or multilingually within these communities will likely be exposed to still more variability; accented speech in their home language, foreign-accented speech and accented speech in their host community language. Being able to deal with variation is crucial to communication, yet relatively little is known about how children acquire sociolinguistic competence, that is how the awareness and use of social-indexical information in speech processing develops. Still less is known about how this might be affected by language background. Previous work has demonstrated that monolingual children aged 5-6 yrs old are able to distinguish their own accent from foreign accented speech, but not their own accent from a different regional accent (Wagner et al., 2014; Floccia et al., 2009, Girard et al., 2008). The current study extends this work to investigate whether accent processing is affected by language background.

Using a similar methodology to that developed by Floccia et al. (2009; see also Wagner et al., 2014), 30 monolingual and 30 bi-/multilingual children aged 5-6yrs were tested in their ability to comprehend and categorize talkers produced in the accent of their local community (London), an unfamiliar regional accent (Yorkshire) and an unfamiliar foreign accent (Singaporean). All children completed a standardized vocabulary test (BPVS; Dunn & Dunn, 2009) followed by a comprehension task, in which they heard either an unfamiliar regional accent or an unfamiliar foreign accent. As well as testing comprehension, this task also exposed children to the unfamiliar accent. The children then completed a speaker categorization task in which they heard stimuli from the same accent they had heard in the comprehension task (i.e., unfamiliar regional or unfamiliar foreign) contrasted with the other unfamiliar accent or their local accent (London), produced by previously heard and novel talkers. Children who were in the condition that did not involve their local accent (i.e., the unfamiliar regional vs. unfamiliar foreign accent) were exposed to only one of these accents in the comprehension task, and the choice of accent was counterbalanced across participants.

All children demonstrated above chance performance in the comprehension task, but language background significantly affected children's ability to categorize talkers of different accents. Monolingual and multilingual children were able to discriminate between their local accent and foreign-accented speech but although monolingual children were able to discriminate between home vs. regional accented speakers with greater accuracy than in previous studies (e.g., Wagner et al., 2014; Girard et al., 2009), they were significantly poorer than multilinguals at categorizing home vs. regional accented speakers. Overall, the results suggest that early exposure to variation in the language environment leads to differences in the processing of sociolinguistic variation.
Grammatical and sociolinguistic constraints on complementizer expression in formal discourse setting

This study uses multivariate analysis to investigate style shifting, linguistic constraints, and the consequences of sentence processing effects on probabilistic distribution of the English complementizer, *that*. Variationists describe *that* in terms of stylistic and language-internal production constraints (Kroch & Small, 1979; Rohdenburg, 1996; Torres & Walker, 2009). English complementizers vary between overt and null forms: “We know CP[*that*/Ø they’re there because of their influence on other objects].”

We explore the role of style and syntax in distribution of *that* by examining a collection of transcribed university lectures given by professors at Yale University ([http://oyc.yale.edu/](http://oyc.yale.edu)), representing a formal discourse setting. We also contrast complementizer variation across discourse registers by comparing our results with those from sociolinguistic interviews, drawn from the Buckeye corpus. We analyzed a sample of 5 courses from disciplines of Psychology, Astronomy, Biology, Religious Studies and History. Nearly 600 tokens were extracted from 11 lectures (2-3 lectures per subject), including all tokens in contexts where a complementizer is licensed. In this formal register, the overall deletion rate was 39%. By comparison, the rate of complementizer absence in the sociolinguistic interviews exceeded 75% of all tokens (N>1200, from 16 speakers). In addition to a difference in the rate of *that* expression across the two formality registers, we also found that lecture type had a significant effect. The highest frequency of zero forms were found in Psychology lectures (57%), compared to religious studies lectures (19%).

We contend that “grammatical ideology,” proposed by Kroch and Small (1978), accounts for this disparity in complementizer expression between the discourse types. K&S argue that prescriptive notions of linguistic correctness favor use of explicit grammatical and formal structures in higher status speakers. Findings suggest that disciplines constitute distinct discursive genres, associated with inherently different levels of prescriptive correctness and formality.

A binomial step-up/step-down analysis using GoldVarb shows that linguistic factors (adjacency of the matrix and embedded verbs, grammatical subject, and main verb lexical identity) significantly affected the variant choice. These results suggest the Complexity Principle also affects the use of *that*. This theory argues that “any elements capable of delaying the processing…of the overall sentence structure favor the use of an explicit signal of subordination” (Rohdenburg 1996, pp. 161). In our data, an increase in verb form and main clause syntactic complexity (e.g., non-adjacency of verb and complementizer) promote *that*.

Additionally, we found significant effects of matrix verb etymology, reflecting Rohdenburg’s complexity principle and Kroch and Small’s prescriptivism. Verbs of Latinate or French origin favored retention of *that*. Rohdenburg proposes that less frequent verbs (like most Latinate verbs in English) trigger increased burden on cognitive recognition of semantic properties, while Kroch and Small argue Latinate verbs are perceived as more prestigious than frequently occurring Germanic verbs, triggering the more prescriptively correct structure.

These preliminary results contribute to the continued pursuit of a nuanced understanding of how formal discourse domain affects complementizer variation. Further analyses will better outline the effects that audience and discourse topic have on a speaker's choice in employing sociolinguistically conditioned syntactic structures.
References


The Dynamic Interaction of Frequency and Internal Constraints: A Case Study of (ING)

Though Pierrehumbert’s (2002; 2006) model of exemplar dynamics explicitly emphasizes the importance of context in the storage and activation of exemplars, relatively little work has tested these claims within the context of sociolinguistic variation (Guy, Hay, and Walker 2008). The question of interaction between internal constraints and frequency bears revisiting, however, since we now have more robust methods in the sociolinguistic toolbox and larger samples with which to tease apart the complex relationship between sociolinguistic variation and word frequency. These tools also allow us to reexamine (ING) with more detail, despite earlier null findings for frequency (Abramowicz 2007). In order to identify how internal constraints interact with lexical frequency, this study analyzes the effect of frequency on (ING), taking into account a word’s frequent context of occurrence.

The data consists of 12,052 tokens of (ING) from 121 speakers, with each token impressionistically coded as either –in or –ing. In addition to realization, tokens were coded for traditional grammatical conditioning factors (Houston 1985) and phonological environment. Frequency was calculated from the SUBTLEX-US corpus (Brysbaert and New 2009) and logarithmically transformed for inclusion in statistical models. To capture a word’s frequency of occurrence in particular environments, a percentage measure for “High-IN” grammatical environments and “Low-IN” phonological environments was calculated from token contexts in the corpus under study. Essentially, if a word occurred in High-IN (i.e. verbal) grammatical environments 80% of the time in the corpus, the measure would capture this bias in grammatical environment in addition to the effects of raw lexical frequency.

Results from mixed-effect logistic regression show a strong effect of frequency on the realization of (ING) in all models fit to the data. More frequent words show a higher likelihood of having an -in realization, suggesting a leniting process at work (Dinkin 2008) that affects the realization of (ING). This frequency effect additionally interacts with grammatical context of occurrence. Frequent occurrence in environments that favor –in amplify the effect of lexical frequency; conversely, environments that favor –ing dampen the effect of overall frequency. The importance of grammatical context of occurrence points to an entrenchment process, where speakers’ sensitivity to heard tokens of (ING) affects their own choice of variant. Furthermore, individual differences in the effect of phonological environment point to a stronger case for (ING) as a morphological, rather than phonological, process (Houston 1985; Tamminga 2014). Overall, these findings support the usage-based position of frequency effects as the result of a dynamic interplay between traditional grammatical constraints and cognitive systems.


Processing variable structures: Object clitics in native and L2 Spanish

The study of variable structures contributes to our understanding of how linguistic forms carry grammatical, lexical, social, and geographic information (e.g. in English particular vowels can indicate geographic origin). Sociolinguistic variation research relies on production data, leaving open questions about the processing of forms that vary according to the geographic region in these occur. Crucially, processing studies have shown that both linguistic and non-linguistic information shapes expectations.[1]

We extend this emerging body of research by investigating the processing of variable Spanish object pronouns. Our study focuses on how language variety and gender influences processing variable structures: the use of dative clitics in accusative contexts (leísmo) in Spanish. Leísmo is geographically determined (Spain), but gender influences use (leísmo often used with masculine referents.[3])

Method: Self-paced reading task investigating whether (non)leísta L2 learners (n=52 advanced, int., beginners) are sensitive to the same factors predicting object-clitic choices as non-leísta natives (n=10). Conditions (2x2): gender (masc./fem.) and case (dat./acc.). Predictions (Tbl. 1): Natives: Processing variants outside of one’s own increases difficulty.[2] Thus, we expect 1) non-leísta speakers will read dative (“le”) clitics more slowly than accusative (“lo/la”) clitics; 2) gender effects: dative clitics + fem. referents will elicit slower RTs. L2ers: previous work shows: 1) preference for masculine forms, 2) gender not a significant predictor, and 3) dative clitics interpreted as simply [+human]. Thus, we predict non-leísta learners might have trouble processing feminine referents with dative and accusative clitics.

Results (Tbl. 2): Two-way repeated-measures ANOVA with case (accusative/dative) and gender (masculine/feminine) as within-participant variables and group as a between-participants variable revealed a significant main gender effect (F(1, 60)=5.808, p<0.05), and a significant three-way interaction (case*gender*group; F(3,60)=2.794, p=.048). As predicted, non-leísta natives 1)read dative clitics more slowly, 2)especially so with feminine referents. Intermediate learners read dative and accusative clitics with feminine referents slower than masculine, confirming developmental trends. Advanced learners, displayed a split pattern, reading dative clitics more slowly only with masculine referents (unlike natives). Like intermediates, they also read feminine accusative clitics more slowly than their dative counterparts (again, unlike natives). These results suggest that L2 learners display sensitivity to only some linguistic factors but not the non-linguistic factors that influence NS-choices in variable structures.

Table 1: Predictions for RTs (region of interest) in all conditions: natives and learners

<table>
<thead>
<tr>
<th></th>
<th>NS</th>
<th>L2ers (lower prof.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: acc.+fem ref.</td>
<td>Low RTs</td>
<td>Highest RTs</td>
</tr>
<tr>
<td>B: acc.+masc. ref.</td>
<td>Lowest RTs</td>
<td>Lower RTS</td>
</tr>
<tr>
<td>C: dat.+fem ref.</td>
<td>Highest RTs</td>
<td>Higher RTs</td>
</tr>
<tr>
<td>D: dat.+masc. ref</td>
<td>High RTs</td>
<td>Lower RTs</td>
</tr>
</tbody>
</table>

Table 2: Mean raw RTs at segment five (clitic-verb) in all conditions for all groups.

<table>
<thead>
<tr>
<th></th>
<th>NS</th>
<th>ADV</th>
<th>INT</th>
<th>INT-Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: acc.+fem ref.</td>
<td>385.87</td>
<td>(80.2)</td>
<td>641.96</td>
<td>(434.1)</td>
</tr>
<tr>
<td>B: acc.+masc. ref.</td>
<td>381.77</td>
<td>(81.1)</td>
<td>581.98</td>
<td>(343.5)</td>
</tr>
<tr>
<td>C: dat.+fem ref.</td>
<td>537.93</td>
<td>(134)</td>
<td>582.42</td>
<td>(309.3)</td>
</tr>
<tr>
<td>D: dat.+masc. ref</td>
<td>429.17</td>
<td>(84.2)</td>
<td>666.79</td>
<td>(349.8)</td>
</tr>
</tbody>
</table>
References


Testing the lexical malleability of a priming effect in Portuguese spoken corpus data

Previous research on priming, where one use of a linguistic variant increases the probability of a subsequent use, in corpus data has found that lexical identity between the prime and the target increases the strength of the general priming effect (Gries 2005). The purpose of the present study is to test the extent to which exceptional items can prime non-exceptional items and vice versa.

Recent studies of the reduction of Portuguese \textit{para} ‘to, for, in order to’ to \textit{p(r)a} have looked at the effect bigram frequency involving the preceding and following words (Gradoville 2012, 2013; Huback 2012). The results of these studies have shown that as bigram frequency increases the probability of reduction also increases, consistent with the reducing effect of frequency (Bybee 2006). Bigrams involving following definite or indefinite articles and the bigram \textit{para que} ‘so that’ were exceptions to this trend in that, despite their relatively high frequency, they had reduction rates far below that found in the rest of the data. This suggests that these exceptional bigrams may be stored somewhat separately in memory and any priming effect where one of the prime or the target is in an exceptional grouping should be weakened. This study aims to test that hypothesis. Given the bigram frequency effect, this study also intends to test whether the conserving effect of frequency (Bybee 2006) attenuates the priming effect.

The data used in this study come from the corpus Português Oral Culto de Fortaleza (Monteiro 1993). The response variable is whether the variant was unreduced \textit{para} or the reduced variant \textit{p(r)a}. All speakers that failed to use both variants in sequence were excluded from the study. The predictors tested in this study include previous occurrence (unreduced \textit{para}, reduced \textit{p(ra)}, non-rhotic \textit{pa}), \textit{para} + WORD bigram frequency (modeled as scalar), target following grammatical category (definite article, indefinite article, \textit{que}, other), prime following grammatical category (same values as previous), following sound (vowel, coronal, dorsal, non-lingual), and following stress (stressed, unstressed). A logistic mixed-effects regression with a varying intercept for speaker was fitted to model the variation under study within R (R Core Team 2015). A three-way interaction term including previous occurrence and the two following grammatical category predictors was used to test the effect of exceptional groupings on priming. A two-way interaction term involving previous occurrence and bigram frequency was used to test the conserving effect of frequency on priming.

The results of this study replicated the previously identified significant main effects for bigram frequency and the following grammatical category. The results also indicate that there is a significant general priming effect as a previous reduced realization (\textit{p(r)a} or \textit{pa}) significantly increases the likelihood of a subsequent reduced variant. On the other hand, neither the three-way nor the two-way interactions yielded a significant result. In other words, the three main effects on this variable are independent of each another. The priming effect, therefore, operates at the same strength regardless of bigram identity between the target and the prime.
References


Effects of talker dialect on lexical decision involving a merged phoneme

Regional language variation poses a problem for speech perception. How is the difficulty of processing speech in an unfamiliar dialect (Floccia et al. 2006, Clarke and Garrett 2004) overcome? Speakers store and use talker-specific information in speech processing (Pisoni 1997; Nygaard et al. 1994) and adjust to a novel feature during an experiment (Maye et al. 2008). Speakers can recognize and label dialects (Clopper and Pisoni 2004a, b). This paper provides experimental evidence that speakers use dialect cues to aid in lexical processing when perceiving speech from unfamiliar talkers.

One potential problem for listeners is differences in phonemic contrasts, e.g., the listener’s variety has two phonemes but another only one. We asked three questions about this aspect of cross-dialect adaptation:

1) What is the lexical status of a word produced with a phoneme replacement consistent with a dialect that is not the listener’s?

In Spain, /s/ and /θ/ contrast, but in Latin America they are merged to /s/. What happens when an unmerged (Castilian) listener hears a word like cine /θine/ ‘movie theater’ pronounced [sine]? Is it processed as if it were a pseudoword, or is it recognized because of exposure to other varieties? 10 Castilian listeners performed single-word auditory lexical decision on words with a merged pronunciation (with /θ/ and /s/ pronounced with [s]). Real word fillers and pseudowords were added for a total of 259 stimuli, heard twice. As expected, the substitution produced a slowdown. Subjects were slower respond to /θ/-type words than to other stimuli (β= 0.166, p < .01, see Figure 1). They do not appear to be pseudowords, but they are not processed like words.

2) Is it easier to process a phoneme substitution when the talker’s accent matches the substitution? That is, is a /θ/-type/ word with [s] processed more easily when produced by a Latin American talker?

To answer this question, the task had two talker conditions, separated by blocks: a Latin American Spanish talker and a Castilian Spanish talker. Each block was preceded by a talker familiarization passage intended to induce dialect recognition. All test stimuli were produced with [s] by both talkers, matching the Latin American talker’s native and presumably expected pronunciation but not the Castilian talker’s. Participants responded to /θ/-type words significantly faster in the Latin American condition (β= -.307, p < .01), suggesting differential expectations for pronunciation of these words by talker.

3) Do all listeners use dialect information in the same way?

This appears to not be the case (Figures 2 and 3). Some subjects rejected or accepted θ-type words in both conditions, but others responded based on on talker accent. This did not straightforwardly relate to differences in response times (Figure 3).

The results show that speakers can extract talker dialect information without overt dialect labels and use it in the processing of speech with regionally variable characteristics. Expectations about talker pronunciation are integrated into the process of lexical access, highlighting a need for speech perception models that incorporate talker dialect information.

References


Figure 1: Comparison of reaction time by stimulus type and talker condition.

Figure 2: Lexical decisions on /θ/-words for each subject by talker condition; accuracy is proportion acceptance. The left shows the responses of those who adopted a different response strategy based on talker dialect, and the right shows those that did not.

Figure 3: Mean reaction time to /θ/-type stimuli for each subject by talker condition.
Word-specificity in Language Variation and Change: Evidence from New Zealand English.

This talk outlines a range of studies conducted on the Origins of New Zealand English (ONZE) corpus, focussing on the degree to which they jointly provide evidence for a role for word-specificity in language variation and change. I argue that, via a production/perception feedback loop, word representations come to be influenced by the distribution of contexts in which we encounter them. Three lines of evidence are presented. First, I examine how word durations change over time, in response to changing distributions of usage factors (such as changes in frequency, informativity, and position in utterance). This provides evidence for phonetically-detailed word-level representations. Second, I briefly examine the much-discussed claim that word frequency is involved in sound change – presenting data from two different types of sound change, that show different types of frequency effect. Both effects result from different aspects of the production/perception loop. Finally, I present data from cases where words are not evenly distributed across particular linguistic or social habitats. Word choices vary systematically across social and linguistic contexts, and these distributions are not stable over time. Some words occur more sentence medially, for example. And some occur more sentence finally. Some words are used more by older speakers, and some by younger speakers. I present evidence that the uneven linguistic and social distributions of words across contexts substantially affects their representations. Taken together, the results presented provide substantial evidence in favour of the existence of detailed word-level representations which reflect the rich range of contexts in which words are encountered.
The role of names in the familiar language advantage

Listeners perform best when dealing with a familiar voice or language (e.g., Nygaard & Pisoni, 1998). This is typically called the familiar language advantage (FLA). With voice memory, FLA manifests as listeners being better at remembering voices with familiar languages or accents. For example, American English-speaking listeners are better at identifying American English speakers than Spanish or Spanish-accented English speakers (Thompson, 1987), while Spanish-English bilinguals are equally good at identifying all three accents (Goggin et al., 1991). The proposed mechanism for these results lies in listeners knowing which phonetic cues are talker-specific in a familiar language but not being able to tease apart talker-specific and language- or accent-general phonetic cues in an unfamiliar language (e.g., Winters et al., 2008; Perrachione & Wong, 2007).

While we know that listeners’ biases and expectations about speakers affects intelligibility and perceived accentedness (e.g., Rubin, 1992), it is unknown how listeners’ beliefs affect voice memory. To this end, we conducted a voice memory study with Mandarin-accented English voices and locally accented English voices in conditions which paired voices with stereotypically congruent names (Mandarin-accented English voice as Chen and locally accented English voice as Connor) and stereotypically incongruent names (vice versa). If the FLA effect relies solely on listeners’ ability to parse phonetic information as talker-specific or accent-general, we would not expect name congruency to affect performance.

Listeners were presented with a sentence read by one of the speakers and asked to identify the name of the speaker. The voices were paired with either Romanized traditional Chinese names (e.g., Chen) or stereotypical white names (e.g., Connor). After being familiarized with 5 voices, listeners completed: (1) training, (2) practise quiz (with feedback), and (3) final test with both accent groups. A total of 112 listeners completed the task, 50 who are English-Mandarin bilinguals and 62 with no Mandarin speaking experience. Listeners performed better on the native English voices \([B = 1.05, SE = 0.13, z = 7.9, p < 0.001]\) and Mandarin-English bilinguals performed better on the task overall \([B = 0.13, SE = 0.05, z = 2.5, p < 0.05]\). There was also an interaction of Accent and Name \([B = -0.37, SE = 0.19, z = 2.0, p < 0.05]\), which indicated that incongruent name/accent pairings negatively affected the native-accented voices more than the Mandarin-accented voices.

These results suggest that voice memory does not solely hinge on listeners’ attribution of phonetic cues. Incongruent name/accent pairings negatively affect performance, particularly for the native accent. One interpretation is that non-native sounding names may bias listeners to perceive a native voice as unfamiliar. Another possibility is that stereotypically incongruent names are common for non-native speakers, but not for native speakers, and thus the association is more difficult to learn for native accents. There may also be an effect of experience with incongruent name/accent pairings on performance; our continuing research tests these effects in a population that regularly encounters a variety of names and accents: instructors at a multicultural university.
Influences of regional dialect, gender, and style on speech intelligibility in two listener populations

Regional dialect, talker gender, and speech style have been shown to affect listeners’ ability to identify speech in noise. Specifically, speech produced in a familiar dialect is more intelligible than speech produced in a less familiar dialect (Clopper et al., 2010), the speech of men is less intelligible on average than the speech of women (Bradlow et al., 1996), and speech produced in a clear style is more intelligible than speech produced in a plain lab style (Picheny et al., 1985). However, previous research has typically focused on only one of these three variables, and these conclusions are based on data collected from relatively homogeneous, undergraduate populations. The goals of the current study were twofold—first, we investigated the influences on intelligibility of these three variables not only in isolation, but also in their interactions. Second, we explored their effects on intelligibility in two distinct listener populations, university undergraduates and a more heterogeneous adult population.

Listeners were presented with auditory tokens of individual words in speech-shaped noise and were asked to type the word that they heard. The tokens were extracted from passages of read speech and varied by talker dialect (Midland vs. Northern), talker gender (male vs. female), and speech style (clear vs. plain). A group of 75 undergraduate students who ranged from 18 to 23 years old (M = 20 years), and a group of 44 adult visitors to a science museum who ranged from 18 to 69 years old (M = 34 years) participated. 43 of the undergraduates and 29 of the museum visitors had grown up in the Midland dialect region. 32 of the undergraduates and 15 of the museum visitors had grown up in the Northern dialect region.

Similar results were obtained for both undergraduates and museum visitors. Listeners more accurately identified words produced by Midland talkers than Northern talkers, regardless of listener dialect, and they more accurately identified words produced in a clear style versus a plain style. In addition, the museum visitors identified words produced by female talkers with greater accuracy than words produced by male talkers. These main effects were mediated by two-way interactions for both listener groups. The talker dialect x talker gender interaction revealed that Midland females were more intelligible than Northern females. Similarly, the clear speech produced by female talkers was more intelligible than the plain speech produced by females. However, neither the effect of dialect nor the effect of style was significant for male talkers.

These contrasting effects of talker dialect and speech style across genders may reflect differences between male and female talkers in patterns of phonetic reduction, which can be affected by both style and talker dialect (Bradlow et al., 1996; Byrd, 1994). A second possible explanation for the interactions may come from challenges to listener expectations. Listeners’ processing may be negatively affected when the speech they hear does not match their expectations about how much phonetic reduction or dialect variation is typical of female versus male talkers (Clopper & Bradlow, 2008).

References:


Reverse Linguistic Stereotyping: The Effect of Listener Expectations on Speech Evaluation

The linguistic stereotyping hypothesis holds that even brief samples of speech varieties (e.g., dialects) associated with low prestige groups can cue negative attributions regarding individual speakers. The converse phenomenon is reverse linguistic stereotyping (RLS). In RLS, attributions of a speaker’s group membership cue distorted perceptions of that speaker’s language. Thus, Rubin and colleagues have repeatedly documented that when listeners mistakenly believe they are listening to a nonnative speaker of English (NNS), they report hearing highly accented speech and their listening comprehension significantly declines. Research has shown RLS not only affects native speakers’ attitudes towards nonnative teachers or international teaching assistants but also their performance and retention of information (Kang & Rubin, 2009; Rubin, 1992). At the session, the presenter addresses the effect of listener expectation and biases on their speech evaluation. Two separate approaches are discussed: (1) native speakers of English as listeners and their proclivity to RLS and (2) non-native speakers as listeners and their proclivity to RLS. The first study establishes a procedure for ascertaining a proclivity to RLS for individual listeners and estimating the contribution of linguistic stereotyping to ratings of NNS oral proficiency. The second study investigates whether the native status of an English teacher affects ESL students’ perceptions of the instructor or teaching ability.

Using a matched-guide technique, the RLS measure exposes listeners (90-150) to two speech samples, both of which were actually recorded by the same speaker. One of the speech samples is associated with a fictitious NNS guise, while the other is associated with a fictitious NS guise. After each speech sample, listeners fill out Zahn & Hopper’s (1988) Speech Evaluation Instrument (SEI), plus several other rating scales and tasks. (A distractor task separates the two guises.) RLS scores are speaker ratings for the NS guise minus ratings for the NNS guise or vise versa depending on the research focus of the native speaker status.

Results replicate in RLS scores Zahn and Hopper’s original three-scale structure: Dynamism, Social attractiveness, and Social Superiority. Regression analyses of the first study reveal certain listener background variables that predict their RLS scores. Findings of the second study show significant interaction between listeners’ cultural backgrounds and comprehension scores but no significant results between nationality and listeners’ ratings. These findings help better understand learners’ perceptions of language proficiency and teaching competence of a nonnative teacher in the context of globalization. Overall, these two studies evince considerable variance in the inclination to engage in RLS, although cross-cultural immersion indices account for only a small amount of that variance. While the RLS measurement procedure outlined here requires more demanding administration than mere paper-and-pencil self-reports, it has the advantage of reflecting authentic RLS processes.
What do you expect from an unfamiliar talker?

One of the longest-standing puzzles in speech perception is how listeners cope with the often extreme differences in how individual talkers use acoustic cues to realize their linguistic intentions. A number of solutions have been proposed, including the recent proposal that listeners quickly adapt to unfamiliar talkers by learning the distributions of acoustic cues that they produce (their “accent”).

This can be formalized as a kind of statistical inference, where listeners try to infer which of all possible accents best explains a talker’s speech (Kleinschmidt & Jaeger, 2015). Prior experience helps because it narrows down the range of possibilities that a listener needs to consider (in Bayesian jargon, it provides an informative prior on accents). We test a critical prediction of this view: when an unfamiliar talker’s accent falls outside the range of typical variation across talkers, listeners should adapt only partially. Specifically, listeners’ phonetic classifications should reflect a compromise between listeners’ prior expectations and the actual accent they hear. We also, in doing so, demonstrate a novel technique for measuring listeners’ subjective prior expectations about an unfamiliar talker’s accent.

We use a /b/-/p/ distributional learning paradigm (Clayards, Tanenhaus, Aslin, & Jacobs, 2008), where listeners (n = 138) hear a bimodal distribution over voice onset time (VOT), with a cluster at a low value implicitly corresponding to /b/ and another at a high value corresponding to /p/. By varying the location of these clusters, we create accents that are more or less like those produced by a typical American English talker (as measured by, e.g., Kronrod, Coppess, & Feldman, 2012) (Figure 1).

We measure how well listeners learn these accents by comparing their classification functions to the ideal boundaries implied by the exposure distributions alone (Figure 2). As predicted, when the VOT clusters were unusually high or low, listeners actual category boundaries reflected a compromise between the typical and exposure talkers.

Second, we used a belief-updating model to work backwards from the patterns of adaptation to different accents, inferring what listeners’ starting beliefs were, and how confident they were in those beliefs. The inferred prior expectations matched the range of typical American English talkers’ /b/ and /p/ distributions, including the counterintuitive finding that listeners were more uncertain about the /b/ mean VOT than /p/, corresponding to the fact that there’s high variance in the VOT of /b/ across talkers due to some talkers prevoicing (Lisker & Abramson, 1964).

The ability to measure listeners’ prior expectations potentially provides an important and heretofore missing tool in the sociophonetics toolbox: it directly links the variability in production of linguistic variables with listeners’ subjective expectations about those variables, both conditioned on social variables. Our proof-of-concept here (implicitly) uses standard American English, but the same procedure can be applied to specific variables like gender, region, class, etc., by providing information to the listener about who the talker is (which listeners do use to guide speech perception, Hay & Drager, 2010; Niedzielski, 1999; Strand & Johnson, 1996).

Figure 1: VOT distributions for each accent.
Figure 2: Listeners’ responses, compared to typical and exposure talkers

References


Teasing apart social and cognitive factors in sociolinguistic perception

That stereotypes affect sociolinguistic processing is by now uncontroversial. Research over the past fifteen years has demonstrated that listeners are attuned to a variety of socio-indexical cues both within and external to the speech signal, and that these cues influence how linguistic features are perceived and evaluated (e.g., Strand 2000; Hay, Warren & Drager 2006; Hay & Drager 2010; Campbell-Kibler 2011). In this talk, I consider how this attested stereotype effect interacts with broader psychological principles governing person perception, including selective attention and category activation/inhibition. I review some recent research on stereotypes in sociolinguistics that has looked specifically at how listeners respond to stereotypically incompatible cues in the speech signal (e.g., Campbell-Kibler 2009; Pharao et al. 2014; Levon 2014). I argue that the results of these studies point to two equally plausible interpretations: one in which social stereotypes serve to “block” the emergence of an incompatible category, and the other in which this apparent “blocking” is an incidental consequence of broader economical processing constraints.

I go on to describe a negative priming experiment (Neill 1977; Tipper 1985) designed to distinguish between these possible interpretations. The experiment is comprised of two related tasks: a priming task and a probe task. In the priming task, listeners rate a speaker for a given social characteristic (age, sexuality, social class) in the course of a standard matched-guise test. The purpose of this task is to prime respondents for the relevant social category. The subsequent probe task is then an adaptation of the Generalized Phoneme Monitoring (GPM) paradigm (Frauenfelder & Segui 1989), in which listeners identify occurrences of TH-fronting (i.e., the labiodental realisation of the interdental fricatives in English) in the same extracts used in the matched-guise test. Given the strong stereotypical association between TH-fronting and working-class speech in Britain (e.g., Kerswill 2003; Stuart-Smith & Timmins 2006), the central question investigated is the extent to which having been primed for a given category in the first task (e.g., “working class” or “gay”) facilitates or inhibits variant monitoring in the probe task. If facilitation/inhibition is found, this would provide strong evidence in favour of the active lateral inhibition of stereotypically-inconsistent traits in perception (e.g., Macrae, Bodenhausen & Milne 1995; Dijksterhuis & van Knippenberg 1996) and thus the notion that social stereotypes can “block” the emergence of incompatible categories, even in the presence of relevant linguistic cues. In the talk, I present current findings from this ongoing work, and discuss the ramifications that these findings have for our understanding of the relationship between social and cognitive factors in sociolinguistic perception.
Listener differences in non-traditional anaphora resolution

Despite being closed class and functional, English pronouns have a rich history of being a site for conscious language change (Bodine 1975). Recent social movements away from the traditional gender binary have again brought pronouns into public discussion, with increasing acknowledgement of non-standard Preferred Gender Pronouns (PGPs): mainstream sites like Facebook now allow users to choose whether they would prefer to be referenced with he/she/they. Since there is evidence that gender is grammatically encoded in pronouns (i.e., Osterhout, Berklin & McLaughlin 1997), we use a speeded grammaticality task to investigate changes in how listeners process non-traditional anaphora resolution.

This study took place in Virginia, and in the recruitment process we explicitly targeted participants of standard college-age along a wide spectrum of different gender and sexual identities. Stimuli for the study were recorded by a young female speaker, and consisted of 60 target sentences, where an unambiguously gendered name was paired with one of three anaphors:

1. Rebecca gave herself a haircut. ("matching")
2. Rebecca gave himself a haircut. ("mismatching")
3. Rebecca gave themself a haircut. ("they")

Listeners were played the target sentences and 80 filler sentences (40 grammatical, 40 ungrammatical) and asked to rate, as quickly as possible, whether they found the sentence unacceptable or acceptable according to personal grammatical standards. Following the GJ task, participants completed a questionnaire that covered their gender and sexual identities and their familiarity with PGPs in everyday life, then were recorded discussing their experimental strategies and their exposure and attitudes towards non-standard PGPs.

30 people have participated in this study so far. Initial analysis shows that compared to a 97.5% acceptance rate of grammatical fillers and matching anaphor constructions, and a 8.3% acceptance rate of ungrammatical fillers, mismatching anaphors were accepted 70% of the time, and they anaphors were accepted only 40% of the time. We ran a logistic regression model with the binary grammatical judgment as the dependent variable, and tested the following independent factors: type of pronoun, trial number, and the social factors covered in the survey and interview. We found a significant interaction between pronoun type and people who recognized the initialism "PGP" (p=0.003). This group, who we interpret as being participants who were highly aware of PGPs and discourse surrounding pronoun diversity, were more accepting of the singular they construction.

The current results show a high degree of acceptance by participants of non-traditional anaphor forms. This could be further evidence that pronouns can be shallowly-processed (see Love 2011). However, the fact that we found differences depending on a participants' awareness of non-traditional pronoun preferences suggests that this acceptability can also be attributed to experience. Moving forward, we will be investigating whether variability in acceptability can be attributed to a change in progress, or if it reflects the specific values of a particular group.
Processing and cross-linguistic influence: An ERP study of second-generation bilinguals and wh-islands

We investigate heritage speakers of Spanish in the US (second-generation bilinguals, SGB), asking how first-learnt languages are processed in populations who are dominant in a later-learnt language (English, in this case). Heritage speakers when compared to native speakers of a Latin American or Iberian baseline are often characterized as having an incomplete, unstable, and/or defective variety of Spanish. By using novel comparison groups and methodologies, we compare SGBs to first-generation bilinguals (FGB), i.e., speakers raised in a Spanish-speaking country before immigrating to the US as adults, and investigate whether the grammars of second-generation bilinguals might be different from other varieties, rather than incomplete.

We use implicit and explicit measures that avoid reading and language manipulation tasks unfamiliar to SGBs. Instead, participants passively listen to auditorily-presented Spanish island constraint stimuli first in an ERP task, which does not require conscious manipulation of the language, and later in an acceptability rating task. In addition to implicit electrophysiological and explicit behavioral measures, participants are administered an auditorily-presented Spanish syntax proficiency task and an oral sociolinguistic survey to probe demographics and language use/background.

We examine the following subtypes of island constraints: wh-movement in sentences with (a) complex noun phrase complements (CC), (b) wh–islands (WH), (c) sentential complements (TT, or COMP-trace), and (d) relative clause islands (RC). These subtypes vary in syntactic structure, reported strength of boundedness, crosslinguistic instantiation, and overlap with English.

In spite of SGB participants performing at ceiling on the proficiency measure, preliminary analyses of the ERP and AJ data show that different stimulus conditions elicit differential responses for SGB and FGB. Additional analyses using mixed-effects modeling will test whether these differences are better predicted by factors related to acquisition, input, and current usage of Spanish than by generational group. These factors include: age of exposure to the second language, English; factors related to socioeconomic status (SES); and patterns of current Spanish usage, include where Spanish is used, with who, and for what.

This study, thus, proposes a fresh approach to the analyses of the grammars of SGBs, situated in a framework of intergenerational change. We argue that in the context of immigrant bilingualism SGB grammars might simply constitute newly changed yet fully systematic grammars, whose outputs happen not to conform to the normative productions with which they are being compared (Cabo & Rothman, 2012; Carreira & Potowski, 2011; Otheguy & Zentella, 2012, pp. 200–215; Putnam & Sánchez, 2013; Rothman, 2007). The stance examined experimentally in this study is that the language of second-generation bilinguals may reflect normal cross-generational lexical and grammatical change, particularly deepened and accelerated by language contact in conditions of minority bilingualism. These changes lead to a native grammar that is different from the monolingual Latin American or Iberian baseline but that is nevertheless complete and systematic.

References:
Effects of message style on responses in texting: Abbreviations, Slang, Spanglish, and Emoticons

While the media has been echoing public concern that text messaging is destroying the language, researchers are exploring the ways in which this “written speech” may really be changing the way we use language (Crystal, 2008). Of particular interest is the effect this technology may have in a world with an increasingly multilingual population. Will code-switching and languages, such as Spanglish, that combine features of two languages (Stevans, 2004) become increasingly common and acceptable? Studies described here explored the effect that style of message has on responses in a multilingual context.

Twenty-five participants, mainly native English speakers, were asked to respond on a mobile phone to twelve messages seen on a computer screen as if they had just received them. Messages were in a) “proper English”, b) “proper Spanish”, c) English with textisms and abbreviations, or d) Spanglish (see examples below). In a follow up study, responses were recorded online.

There was a significant effect of message style on response; in forty percent of cases, the participant matched the style of their response to the style of the original message. Twenty-two of the 25 times Spanglish was used by participants was in response to a message in Spanglish, and similarly with Spanish. There were 47 examples of matching in the “proper English” condition, and overall most responses (151) were in reasonably correct English. Ninety-eight responses were in English with abbreviations and/or textisms.

Some abbreviations or textisms were widely used. Frequently, letters were used phonetically; there were 9 uses of u for you by five different people, yu was used twice, and ur once. K was used twice for okay, Kay once, y for why once, and n for in was used twice. Using numbers as phonemes was also common, with two examples of 2 and 4, used singly as well 2night, 2mrw, 2day, 2marrow, 4gdt, 4got, and b4. (Coding responses to this type of material was sometimes challenging.)

There were striking individual differences primarily in the use of textisms. Of the ten uses of emoticons, nine came from one individual. Of four uses of @, three were from one person.

Most participants knew some Spanish, but only four considered themselves fluent. Seven participants used Spanish or Spanglish appropriately, while another five used some Spanish, including one who responded to Spanish messages with No hablo espanol. Other examples were: “Hola chica!!”, “No habla espanol!” And in Spanglish: “Yeah Im down for la playa.” and “Translate for me por favor?”

In a third experiment, 60 participants were asked to respond to text messages on a computer as if received on their phone. In this case, the “person” sending the message was identified; BFF, school mate, TA, Boss, and “Mom,” and each had a different style of message. Again the responses showed effects of message style. Matching occurred most commonly for formal English.

Currently we are working on the effects of message style on reading times, and of emojis on comprehension.
Stimuli used:

a) Hey, I haven't heard from you in years!! How is life treating you?

b) Vamonos para la playa este fin de semana, o se te ocurre algo más?
   *(Let's go to the beach this weekend, unless something else occurs to you?)*

c) I'm gonna b @ Lori's house in da afternoon do u want 2 stop by & hng out w us?
   *(I am going to be at Lori’s house this afternoon. Do you want to stop by and hang out with us?)*

d) Hey, y is da casa a mess? Did u 4get about la fiesta 2nite?
   *(Hey! Why is the house a mess? Did you forget about the party tonight?)*

References


Successful Referential Learning in 14-Month-Olds with Minimal Pair Labels using Social Cues and Eye Tracking

Infants depend on a variety of cues to learn labels and referents, allocating attention across general, linguistic, and social domains. Typically, a caretaker presents an object and labels it as the infant perceives the caretaker’s intention, object properties, and label properties. Currently we know little about how these domains integrate with each other, guiding attention under various challenging situations. This study investigated whether ostensive cues, particularly the speaker’s visual regard, assist infants when associative word-object learning is difficult. Capitalizing on previous work showing diminished associative learning between novel objects and novel labels when the labels are highly similar (i.e., minimal pairs), this study explored whether gaze reference augmented infants’ ability to learn minimal-pair object-label relations.

Twenty-two 14-mo-olds (M=14.7; SD=1.12; 9 females) were familiarized to two object-label pairs (objectA+labelA, objectB+labelB; e.g., “din” and “bin”) presented on a Tobii T60 eye tracker. Four familiarization trials (30 s each) consisted of a woman’s face+voice looking forward toward the infant then turned toward the object. When facing forward, she labeled the object in an infant-directed style 8 times/trial. After familiarization, testing consisted of two “same” (e.g., objectA+labelA) and two “switch” (e.g., objectA+labelB) trials. Due to subject loss in the second set, only the first two test trials are presented here (same1; switch1). For analysis, areas of interest (AOIs) were drawn on objects, eye region, and mouth region separately for forward looking v. turned toward object.

A 2x2x3 ANOVA was conducted on trial type (switch, same), location (forward, turned), and AOI (object, eyes, mouth) on proportion of looking time. First, a main effect for trial type was found with greater looking on the switch than same trial (F(1, 20)=9.90, p<.01, η²=0.01, η_g²=0.32; M_{Switch}=0.26; M_{Same}=0.22). Second, a main effect for location was found with greater looking when the speaker was looking forward compared to when she was turned (F(1, 20)=5.07, p=.04, η²=0.01, η_g²=0.19; M_{Forward}=0.25; M_{Turned}=0.23). Third, a main effect for AOI was found with greater looking at the mouth than any other region (F(2, 19)=69.90, p<.01, η²=0.80, η_g²=0.88; M_{Object}=0.16; M_{Eyes}=0.03; M_{Mouth}=0.54). Last, a significant Trial Type x Location x AOI interaction was found (F(2, 19)=4.63, p=.02, η²=0.01, η_g²=0.32). During both switch and same trials, infants looked significantly more at the speaker’s mouth than any other AOI, especially when she was looking forward. However, infants looked significantly more at the speaker’s mouth when she was looking forward during the switch than the same test trial.

Taken together, these results show successful word learning with minimal pair labels. The availability of ostensive cues (gaze regard; infant directed speech) abated the cognitive load that seems to compromise infants’ minimal pair associative learning (at least within this age range). Interestingly, 96% of the variance in this model was accounted for by AOI (main effects and interactions); specifically, differential attention to the speaker’s mouth overall but particularly during the switch trial emerged as a key factor of importance. That is, infants seem to be using the mouth as a visual means of disambiguating the incongruent word/object pairing.
**Proportion of Looking Time as a Function of Trial Type, Location and AOI**

<table>
<thead>
<tr>
<th></th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>Partial eta</th>
<th>Classic eta</th>
</tr>
</thead>
<tbody>
<tr>
<td>TrialType</td>
<td>0.097</td>
<td>8.85</td>
<td>0.01</td>
<td>0.31</td>
<td>0.01</td>
</tr>
<tr>
<td>TT error</td>
<td>0.011</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>0.054</td>
<td>10.10</td>
<td>0.01</td>
<td>0.33</td>
<td>0.01</td>
</tr>
<tr>
<td>L error</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOI</td>
<td>6.207</td>
<td>87.244</td>
<td>0.001</td>
<td>0.81</td>
<td>0.80</td>
</tr>
<tr>
<td>A error</td>
<td>0.071</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T x L</td>
<td>0.001</td>
<td>0.24</td>
<td>0.63</td>
<td>0.01</td>
<td>0.0001</td>
</tr>
<tr>
<td>T x L error</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T x A</td>
<td>0.165</td>
<td>6.54</td>
<td>0.003</td>
<td>0.25</td>
<td>0.02</td>
</tr>
<tr>
<td>T x A error</td>
<td>0.025</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L x A</td>
<td>1.069</td>
<td>47.52</td>
<td>0.001</td>
<td>0.70</td>
<td>0.14</td>
</tr>
<tr>
<td>L x A error</td>
<td>0.022</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-way</td>
<td>0.052</td>
<td>5.82</td>
<td>0.006</td>
<td>0.23</td>
<td>0.02</td>
</tr>
<tr>
<td>3-way error</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**MEANS**
ME Trial Type: Switch (.26), Same (.22)
ME Location: Forward (.26), Turned (.23)
ME AOI: Object (.15), Eyes (.03), Mouth (.55)

<table>
<thead>
<tr>
<th>Trial x AOI</th>
<th>Switch</th>
<th>Object</th>
<th>Eyes</th>
<th>Mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouth &gt; Object &gt; Eyes</td>
<td>.17</td>
<td>.01</td>
<td>.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Same</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouth &gt; Object &gt; Eyes</td>
<td>.14</td>
<td>.05</td>
<td>.48</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location x AOI</th>
<th>Forward</th>
<th>Object</th>
<th>Eyes</th>
<th>Mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouth &gt; Object = Eyes</td>
<td>.06</td>
<td>.03</td>
<td>.68</td>
<td></td>
</tr>
<tr>
<td>Turned</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouth &gt; Object &gt; Eyes</td>
<td>.24</td>
<td>.03</td>
<td>.42</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Trial Type x Location x AOI</th>
<th>Switch</th>
<th>Forward</th>
<th>Object</th>
<th>Eyes</th>
<th>Mouth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouth &gt; Object &gt; Eyes</td>
<td>.06</td>
<td>.01</td>
<td>.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turned</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouth &gt; Object &gt; Eyes</td>
<td>.28</td>
<td>.01</td>
<td>.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Same</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouth &gt; Object = Eyes</td>
<td>.06</td>
<td>.06</td>
<td>.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turned</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouth &gt; Object &gt; Eyes</td>
<td>.21</td>
<td>.04</td>
<td>.38</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Exploring social meanings of morphosyntactic variation using the matched guise paradigm

While recent work has explored the social meaning of phonetic variables (Campbell-Kibler 2009; Pharao et al. 2014), the social attributes that morphosyntactic constructions carry has been less studied (though see Bender 2001; Squires 2013). In this study we use a matched guise technique (Lambert et al. 1966) to investigate the social meaning of two variable constructions attested in American English: double modals (“I might could go to the store” - a feature of some Southern dialects) and "needs washed" (a feature of Midland dialects).

Our study was run online. Participants were told that they would hear fifteen different speakers; in fact, they heard eleven different speakers (all young, American females), but four sentences repeated. Two of the repeating sentences contained "might could" constructions, intact in one version, and with the "could" cut from the recording in the other version. The other two repeating sentences contained "needs to be" constructions, in one version intact, and in one version with the "to be" cut. After listening to each sentence, participants were asked to rate how educated, accented, foreign, friendly, feminine and certain (of the proposition) each speaker sounded. Participants were then asked to list where they have lived, their age, their gender, and whether double modal and needs washed sentences were grammatical to them.

Data collection is ongoing, and 111 participants have completed the study so far. We fit different linear regression models to each of the attribute ratings. For the modal constructions, we find that the double modal guise is rated as more southern and less certain than the standard version. It is also rated as less educated, but only by older listeners. For the needs constructions, we find that the "to be" version is rated as more educated overall, and that Southern listeners - who are generally more likely to rate all speakers as sounding Southern - rate the non-standard construction as explicitly not southern.

These results show that morphosyntactic variables carry social meaning, here specifically associations with region and education. However, the results also suggest that different listeners respond differently to the variation - for double modals, older listeners interpret it differently than younger listeners (Hasty 2011; Eckert 2014), and for needs+Ved constructions, southern listeners (accurately) rate it as non-southern, but other listeners do not.
The Emerging Power of Attention to Facilitate and Guide Early Language Learning in Human Infancy

Over the past 10 years, my students and I have examined patterns of infants’ attention to speaking faces using both visual fixation and eye tracking protocols, under a variety of conditions. We have also explored predictive relations between infants’ selective attention to speakers and measures of concurrent and future language skills (e.g., ASQ Communication, MCDI I and II). We have included infant participants from approximately 8- to 20-months of age, and have consistently seen some interesting patterns of selective attention to speakers that are relevant to our understanding of how attention and its allocation foster early language skills.

First, when viewing a speaking female alone, all aged infants attend most to the mouth region, although the second most scanned region is the eyes. Importantly, this pattern is only observable during dynamic presentations of female speakers; when the face is static, no differential attention to the mouth region is evident, even when a voice track is concurrently available. However, this mouth-region bias holds for silent presentations of a moving face but not when a speaking face and voice are incongruent. Interestingly, we have also found a bias toward the mouth region even when the speaker was inverted (i.e., mouth now appears in the typical eye region). Overall, it seems likely that infants are highly sensitive to areas of speakers that are moving, but only in ways that preserve the intersensory integration of voice+visual articulators.

Second, under richer contexts such as when a speaker is shifting their attention from the infant to an object while naming, infants tend to shift attention to the speaker’s eyes if she visually regards the object. However, the amount of attention to the eyes is not predictive of successful object+label learning, but attention to the mouth is. In fact, recent work in our lab has found that infants’ attention to the mouth of a speaker as she labels an object helps them learn this relation even cognitive load is high (e.g., using minimal pair labels).

Third, we have found some evidence for relations between infants’ face-scanning patterns and measures of receptive and productive language skills (e.g., MCDI, Vineland, ASQ). For example, 11-month-old infants who were less distracted by a peripheral visual distracter in the presence of a female speaker had significantly higher expressive vocabulary sizes at 18 months. However, it remains unclear whether infants’ attention to speakers (and more specifically, their mouths as they are speaking) is uniquely predictive of growing language facility. It is possible that infants’ attentional control per se is paramount to emerging language skills, and that this ability is co-opted during dynamic interactional exchanges between infants and others in a productive way. Thus, attention to any one particular region of a speaker’s face may not be as relevant to the process of language learning, as is the ability to flexibly focus attention to salient, rather than irrelevant, information during dynamic exchange. I will present some ideas of how to try and clarify the role of early attention for emerging language competence.
Figure 1. Percent duration for valid fixations (n = 22 infants) for two areas of interest (AOI): eye region and mouth region in two conditions: dynamic female face+voice (infant-directed) compared to static female face+voice (infant-directed).

Figure 2. Average duration for valid fixations (n = 22 infants; different from those in Figure 1) for three areas of interest (AOI): eye region, mouth region, and not AOI (area not including the face) in three conditions: congruent female face+voice (infant-directed), incongruent female face+voice (infant-directed), and silent female face (infant-directed).
Speaker age, listener age and vowel perception

Recent studies have shown that information about speakers can influence vowel perception, particularly the placement of a boundary between adjacent phonemes in the vowel space can be affected by beliefs about the speaker (Niedzielski (1999), Drager (2011), Koops (2008)).

This study concerns differences in the placement of the boundary between the two vowel phonemes /ɛ/ and /a/, which have undergone near merger within the past few generations in Copenhagen Danish with /a/ being raised closer to /ɛ/. The raised variant of /a/ was overtly stigmatized at the beginning of this change in the 1970s through the first half of the 1980s, but today the stigma appears to have been lost.

17 older listeners and 16 younger listeners participated in a forced choice lexical identification task of stimuli using 6 minimal pairs that varied between /ɛ/ and /a/ in a synthesized 8 step continuum. Each participant listened to 4 different voices, 2 older and 2 younger voices, in order to investigate whether speaker age would have an effect on the placement of the boundary.

The results show an effect of listener age where the younger listeners accept a greater range of vowels in the continuum as tokens of /a/ compared to the older listeners. There was no effect of perceived speaker age. A comparison of the production of /a/ by a set of age matched speakers taken from a corpus of sociolinguistic interviews show no difference in speech production between the two generations. It is therefore suggested that the difference in perception cannot be due to differences in production, but instead could be explained by loss of social stigma previously associated with the raised variant of /a/.

References:
Can prosody disambiguate mathematical structure? It depends who's listening.

Speakers use prosodic manipulations including pitch changes, lengthening, and pausing to indicate the intended grouping of words in their utterances [1,2], and listeners use these cues to help determine the syntactic structure the speaker intended [3,4]. While a handful of studies have examined differences in the prosodic cues used by speakers of different dialects [4] and ethnicities [5], few if any have looked at variation in how different groups of listeners make use of these cues. The present study investigates whether listeners' ability to selected the intended referent of a read mathematical utterance varies with age (and, in ongoing work, occupation and mathematical skill).

Mathematical expressions provide an ideal ground to look for variation in prosodic disambiguation. Phrases like (1a) are ambiguous between two or more mathematical structures (1b,c) that unambiguously show the structural relationships between terms of the expressions. While the ambiguity of (1a) is identical to that of everyday English phrases like (2), we expect all subjects to be familiar with (2) while exposure to mathematics can vary widely within a speech community. Comparing performance on prosodic disambiguation of (1a) and (2) allows study of how variable exposure to mathematics influences a listener's ability to apply general rules of intonational phonology to this specialized domain.

(1)  
a. Nine times A plus four  
b. \((9 \cdot A) + 4\)  
c. \(9 \cdot (A + 4)\)

(2) Sam and Steve or Bob will come to the party.

In a perception study, listeners aged 7-59 judged which of two similar mathematical expressions a speaker intended when using a phrase like (1a), or which of two paraphrases best matched the intended meaning of phrases like (2). Mathematical stimuli came from a related production experiment [6], and on critical trials featured prosodic phrasing that used pitch changes and prosodic breaks to group the variable \(A\) with either the preceding or following term. Everyday English trials relied on ambiguities with prepositional phrase attachment, relative clause attachment, or complex NP conjunction, and were likewise disambiguated by prosodic phrasing.

![Effect of Age on Disambiguation](image)

*Figure 1.*
A linear mixed effects analysis showed significant effects of both age and English paraphrase accuracy on the ability to select the intended mathematical expression, and a significant interaction. Figure 1 shows the accuracy in selecting the intended mathematical expression as a function of age. The youngest subjects scored below chance, but a strong correlation with age was observed throughout childhood ($t = 62.42$, $p < 0.001$, $r = .882$). Adult performance was much more varied, with age a significant but weak predictor ($t = 3.2$, $p < 0.005$, $r = .103$). Adult disambiguation accuracy was instead more closely tied to their performance on the everyday English trials ($r = .712$, vs .588 for children).

Variation in adult and child performance is marked. Presumably children gain experience throughout childhood listening to others read math problems, accounting for the strong correlation with age. Some adults continue to use math; others avoid it whenever possible. An ongoing continuation of this study collects more demographic information on math use and attitudes, and includes collegiate subjects.

References

Linking perception and production in sound change

In this paper we report the results of a study on the link between the perception and production of two sound changes in progress in Dutch, at both the regional and the individual level.

The devoicing of word-initial labiodental fricatives and bilabial stops both show patterns of regional variation in the degree of devoicing (Van de Velde et al. 1996, Kissine et al. 2004). These sound changes however differ in their degree of completion (advanced change vs. incipient change). Completed changes might result in a merger of /v/-/f/ and /b/-/p/. Five regions in the Dutch language area were selected to represent different stages of sound change (West-Flanders, Flemish-Brabant, Limburg, South-Holland, Groningen). For each region, 10 men and 10 women, highly educated, and between 18 and 28 years old were selected (k=100) and participated in a series of production and perception experiments.

First, we discuss the results of two forced-choice categorization tasks, in which participants categorized labiodental fricatives and bilabial stops as being either voiced or voiceless (k=405). Fricatives were phonetically manipulated along the dimensions voicing and duration on the continuum from [v] to [f], and stops ([b]-[p]) along the dimensions voice onset time and duration. The data were analyzed with a mixed effect logistic regression. The perception of fricatives was the most categorical in regions where the devoicing process is still in an incipient phase, and the least categorical in regions where the process of devoicing is almost completed. These results showed clear patterns of cross-regional differences in the speech perception of Dutch (like Kendall and Fridland (2012) showed for English). In contrast, the perception of stops yielded only individual differences.

Second, we discuss the production data. Each participant was recorded in a range of different speech styles, with different degrees of monitoring (word reading, sentence reading, semi-spontaneous speech and spontaneous speech). All fricative and stop tokens were measured along the same phonetic dimensions as in the perception tasks. The production patterns were analyzed between regions and compared to the results of previous production studies.

Finally, we investigate the link between the production and perception of variation at the individual level. We observe a clear link between these systems undergoing sound change. When putting the results of fricatives and stops together, it appears that changes in the perceptual system precede changes in production. Language users first seem to incorporate a new form in their perception, and only later in their production.

Reference


Static Human Gaze Activity Does Not Diminish Infants’ Cognitive Load during Referential Learning

Werker et al. (1998) habituated 14-month-olds to two object-label pairings (ObjectA+LabelA and ObjectB+LabelB), and then tested infants’ discrimination of a subsequent mismatch (ObjectA+LabelB; called the ‘switch’). When the labels were maximal pairs (e.g., lif v. neem), infants significantly recovered attention on the switch trial. In contrast, Stager and Werker (1998) found no significant recovery of attention on switch trials with minimal pair labels (bih v. dih), suggesting that the cognitive resource requirement of the task is essential to referential learning. In fact, lessening the cognitive demand (e.g., familiarizing infants with minimal pairs in advance of the task; Fennel & Werker, 2003) augments infants’ ability to learn difficult (e.g., doll and ball) object/word pairings. The purpose of this study was to examine whether the inclusion of a female face, making eye contact with the object being labeled, would also act to decrease cognitive load in a minimal pair referential task.

Using an infant-controlled switch task, a total of 18 13- to 15-month-olds were presented with object-label pairings in one of two conditions: maximal pairs (e.g., jode and pake) or minimal pairs (e.g., bin and din). Importantly, a female face first appeared on the screen, looking directly at the infant, then shifted gaze to the right (or left) in advance of the object+label presentation (on the corresponding side of the gaze shift). Infants’ looking times were measured and compared across test trials and groups, resulting in a significant trial type x group interaction (F(1, 16)=7.12, p=.02, ηp²=0.31). Infants in the maximal pair group looked longer on switch trials (M_{switch}=10.13, SD=1.8) than same trials (M_{same}=6.83, SD=1.03), whereas infants in the minimal pair group did not (M_{switch}=4.76, SD=1.76; M_{same}=6.85, SD=1.46). Importantly, no significant difference in looking times on the objects and faces was found within each condition (Minimal, p=.20; Maximal, p=.37), nor between conditions (Faces, p=.83; Objects, p=.88).

These results replicate the general finding that as cognitive load increases (e.g., minimal pair object labels), infants have more difficulty learning object-label relations. The addition of a female gaze-relevant reference did not decrease this cognitive load, in spite of the value of gaze contingency for other aspects of early language learning (e.g., Senju, Csibra, & Johnson, 2008). It is possible that in order for human gaze-relevancy to attenuate cognitive load in referential tasks, other ostensive forms of communication must also be present (e.g., smiling; the label emanating directly from the agent). It also possible that mutual gaze was not established in the current study given no formal contingency between where infants’ looked before gaze shift was implemented. In this vein, Fais et al. (2012) demonstrated that infants’ success with minimal pairs in referential learning was positively associated with sharing mutual gaze with a live experimenter, suggesting that this may be related to emerging joint attention skills. We will discuss ways in which our own “social switch” protocol can be improved to directly test this hypothesis (e.g., the use of gaze-contingent eye tracking) as well as other ostensive cues.
Picking Apart Perceptual Recalibration: An Exercise in Applying Experimental Methods

In listening to speech, people have been shown to apply several types of adjustment to their phonemic categories that take into account variations in the prevailing linguistic environment. These adjustments include selective adaptation (Eimas & Corbit, 1973), lexically-driven recalibration (Norris, McQueen, & Cutler, 2003), and audiovisually-determined recalibration (Bertelson, Vroomen, & de Gelder, 2003). The current study focuses on lexically-driven recalibration. This phenomenon occurs when listeners hear a number of words that contain a slight mispronunciation. For example, a word like “mentorship” might be presented with a pronunciation that is midway between “mentorship” and “mentorsip”. After hearing number of words with such mispronunciations, listeners adjust their phoneme boundaries, expanding the category so that the formerly ambiguous sounds are now perceived as acceptable tokens.

A recurring psycholinguistic question has been whether these adjustments are an automatic consequence of speech perception, requiring no cognitive resources or whether instead such adjustments rely on some kind of attentional investment. This question has been posed by giving listeners an additional task during their speech perception. For selective adaptation (Samuel & Kat, 1998), audiovisually-driven recalibration (Baart & Vroomen, 2010), and lexically-driven recalibration (Zhang & Samuel, 2014) this approach has produced a very consistent answer: In all three cases, no loss of adjustment has been found under load, supporting the view that these adjustments do not require attentional resources.

In the current study, a “targeted distraction” procedure was used instead of the standard dual-task method. The key difference is that a distracting load is imposed at a precise moment, rather than via some kind of ongoing task. The latter case potentially allows listeners to “time share”, switching attention between tasks strategically. Targeted distraction aims to eliminate this possibility, and with this method the data indicate that lexically-driven recalibration does in fact require attention. Building on this finding, the targeted distraction method is used to measure the period of time during which the lexical percept remains malleable. The results support a processing window of approximately one second, consistent with the results of a small number of prior studies that bear on this question. The method also allows us to examine the type of resources that limit performance (e.g., specifically language-oriented or more general), and the relationship of the recalibration process to lexical activation.

Effects of regional variation on speech processing for second language learners

Previous research has found that in some cases unfamiliar regional and social varieties of the language may have an effect on comprehension, for both native and nonnative listeners (e.g., Eisenstein & Berkowitz, 1981; Major et al., 2005), and that variable sounds may be perceived and processed differently according to the listener’s native variety (e.g., Evans & Iverson, 2004; Sumner & Samuel, 2009). However, increased familiarity with dialectal speech may lead to lessening effects of variation on speech comprehension, even for second language (L2) listeners (Tauroza & Luk, 1997). Furthermore, exposure to specific regional varieties has been found – in some cases – to lead to changes in how some regional sounds are perceived (e.g., Escudero & Boersma, 2004; Baker & Smith, 2010). The current study explores how observed improvements in intelligibility and perceived comprehensibility (Derwing & Munro, 1997) of dialectal speech by L2 listeners may be attributed to changes in how dialectal phones are processed. Specifically, the study investigates the effects of regional and social phonological variation on L2 speech processing in terms of:

(a) lexical activation (Is the intended lexical item activated and/or comprehended?)
(b) processing costs (How does the presence of variable sounds affect processing time?)

A Lexical Decision Task was administered to two groups of listeners: (1) 50 English-speaking university students of Spanish in the U.S. Midwest (experimental group) and (2) 15 native speakers of Spanish in Buenos Aires, Argentina (control group). Listeners heard both Spanish nonce and real words with dialectal and standard phones and were asked to indicate whether each item heard was a Spanish word. The target conditions included (a) the standard full sibilant /s/ (basta [basta]) and dialectal aspirated /s/ (pasta [pahta]), and (b) the palatal fricative [ʝ] (playa [plaja]) and the Argentine assibilated palatal [ʃ] (mayo [maʃo]). Reaction time latencies and word judgment accuracies were analyzed according to: listener group (L2 group, native control), L2 proficiency level (intermediate, advanced), L2 dialect exposure (Argentine Spanish exposure, no exposure), phone (standard [s, j], dialectal [h, ʃ]), and lexical status of the stimulus (real, nonce).

Findings from the data analysis revealed an effect of the dialectal variants on processing the real word (but not the nonce) stimuli for the L2 listener group. Slower L2 reaction times and less accurate lexical decision judgments were observed for (a) items with the dialectal aspirated /s/ as opposed to the full sibilant /s/ variant, as well as for (b) items with the dialectal assibilated palatal variant [ʃ] over items with the more widespread palatal fricative form [ʝ]. Furthermore, while there were no differences across proficiency levels, differences in processing effects within the L2 group were observed according to individual dialect contact experiences. Findings from the study reveal that there may be processing costs for L2 listeners when listening to speech characterized by unfamiliar regional or social/stylistic variants but that those costs may lessen with sufficient exposure (dialect contact).
References


Gender and gendered speech expectations in perceptual adaptation

A variety of studies have shown that there are gender based differences in production of speech sounds that are much greater than the sum of physiological differences between male and female vocal tracts, and that listeners are sensitive to these gendered differences. For example, Li, et al. (2008) showed that Mandarin speaking children developed significant gender differences in the average spectral frequency of their /s/ productions by the age of 5, and Strand and Johnson (1996) showed that American listeners had different category boundaries between /s/ and /ʃ/, depending on whether they believed they were listening to a male or female voice. Less has been studied on gender differences in the alveolar stops, however, we can posit at least a physiological difference, with male productions having a lower COG (center of gravity) than females’ due to vocal tract length. Perceptual adaptation, the act of remapping phoneme boundaries to attune to a particular individual or accent group, has not been studied as much in the context of gender as it has been for known and novel accent types, but it should, hypothetically, be sensitive to gender differences in the case where such differences are encoded in the listener’s phonological system.

A perceptual adjustment experiment was designed to elicit a perceptual split of /t/ before /w/ (towards a front variant, /ʦ/, and a retracted variant /ʧ/), depending on the gender of the talker. Participants who heard only a single variant during training displayed perceptual adjustment, in which the boundaries of the /t/ category (before /w/ as well as before /u/) shifted in the direction of the training variant, and more tokens of that type of variant were acceptable as instances of /t/ than for untrained participants. But participants who heard male voices using the front variant (ʦ), and female voices using the retracted variant (ʧ), showed increased acceptance of /ʦ/ spoken by men, but not by women. They also accepted more /ʧ/ tokens as instances of /t/ for female talkers, relative to the untrained group, but also accepted more /ʧ/ tokens for male talkers. Participants who were trained with women using /ʦ/, and men using /ʧ/, displayed overall perceptual broadening, in which the /t/ category expanded to accept both variants spoken by talkers of either gender as instances of /t/. It is hypothesized that because the men’s retracted productions had the lowest frequency spectral average, and women’s front productions had the highest, the range of variation was greater in this scenario than when the distribution of variants was reversed. The greater range of variation caused perceptual broadening for participants in this condition. When the distribution of variants was along socially expected lines (male retracted, female front), gender may have been disregarded as a relevant factor in adapting to the new pattern, but when the distribution was unexpected (male front, female retracted), the least expected variant (/ʦ/) was sorted according to gender.

References:

**Voices and variation: What we can learn about processing and representations by investigating spoken word recognition and memory**

Episodic theories of representation and lexical access are now strongly supported with a wide range of data. With this in mind, we might wonder why idealized variations of words seem to have a benefit both in terms of spoken word recognition, and detailed memory retention. This pattern runs counter to our hypotheses and theories about the atypical, infrequent linguistic unit, from an episodic perspective. In this talk, I provide an overview of this bias, illuminate instances in which our assumptions about the phonetic composition of a word have contributed to the bias, and show that even after such considerations, these less common pronunciation types do indeed have a processing benefit that does not match with our theoretical expectations. Next, I discuss an approach to experience, which includes the social weighting of linguistic events during perception. In other words, all linguistic events, or instances of spoken words, are not weighted equally by the system. This idea of weighting leads to an explanation of the idealized-form benefit, while building on the foundation established from episodic theories. Finally, I relate this work to research investigating social effects in speech processing more broadly and discuss the importance of claims we can make about individuals versus claims we might want to make about groups from our studies.
Speaker sex and the processing of grammatical gender: Evidence from Italian-English bilinguals

Recent monolingual research in psychology suggests a facilitative effect of congruent speaker sex and grammatical gender during speech perception, such that a feminine word is processed faster when spoken by a female speaker. Vitevitch et al. (2013) report this effect in a Spanish grammatical gender identification task, where participants were faster and more accurate when the speaker’s sex ‘matched’ the grammatical gender of the word (e.g. the feminine word ‘luna’ spoken in a female voice). Andanova (2013) found the same effect in both a grammatical gender identification task and a cued shadowing task using Bulgarian and Italian.

The present project takes this idea a step further, and investigates whether a similar facilitative effect of speaker sex can be observed in a bilingual’s L2 that has no grammatical gender, but where the corresponding L1 word has a congruent grammatical gender. For example, do Italian (L1) – English (L2) bilingual listeners process the English word ‘moon’ faster when spoken by a female speaker, because in their L1 the translation equivalent is a feminine word, ‘luna’?

Based on previously proposed connections between the L1 and L2, the prediction is that this facilitative effect does exist across the two languages. Figure 1 illustrates the conceptual framework and the hypothesized operation of the effect. Hearing the L2 English word ‘moon’ activates the L1 Italian translation equivalent ‘luna’ through the conceptual link proposed by psycholinguists (e.g., Kroll & Stewart 1994). As ‘luna’ is a feminine word it activates “feminine-ness” in the speaker’s mind, connecting to “female voice” through the gender link demonstrated by Vitevitch et al (2013) and Andanova (2013). Finally the socio-indexical link between the two languages, put forward in Szakay et al. (2012), connects “female voice” in the L1 to “female voice” in the L2. Thus, through these connections, the L2 English word ‘moon’ is processed faster when spoken in a female voice, despite the English language itself not having grammatical gender.

Data collection is ongoing at this stage, however, preliminary results from a cued shadowing task by Italian-English bilingual listeners confirm the above predictions, and show a facilitative effect of congruent speaker sex and grammatical gender across the bilingual’s two languages. Listeners are faster and more accurate when the speaker’s sex matches the grammatical gender of the word in their L1. As expected, English monolingual participants do not show the effect, i.e., there is no processing benefit when ‘moon’ is uttered by a female rather than a male speaker.

The results of the study suggest that indexical information can influence higher-level grammatical processing across the L1 and L2, and in general shed more light on the cognitive architecture of language processing. In particular, the study highlights the importance of taking indexical information into account during bilingual speech processing, an area of research that has not received ample attention so far.
REFERENCES


Happy Faces + Voices Facilitate 12-Month-Olds Word-Object Learning

Numerous studies examine emotion and language development during infancy, albeit separately. However, evidence from studies investigating the learning-promoting features of infant-directed speech (IDS) suggests that it may actually be the happy tone typically present in IDS that draws infants’ attention (Trainor, Austin, & Desjardins, 2000). This study examined whether socioemotional cues (e.g. neutral or positively-valenced facial + vocal emotion expression) influenced associative word-object learning in 12-month-old infants. Specifically, is learning enhanced when a speaker is happy vs. neutral?

Using a Tobii© T60 eye-tracker, 20 12-month-old infants were familiarized and tested on four word-object associations that varied in bimodal emotion (happy vs. neutral). Familiarization trials comprised actresses looking towards and labeling a target object in either a happy or neutral tone while ignoring a non-target distractor object on the opposite side of the screen. During test trials, the familiar target object and non-target object were simultaneously presented on the screen with audio prompts (i.e. “Where is the boog?” and “Can you find the boog?”).

Infants were expected to demonstrate better learning of word-object associations during the happy test trials by exhibiting differential attention to the two objects. This hypothesis was supported: planned comparisons indicated that infants had a preference for the non-target object ($M = 37.53\%$, $SE = 1.88$) compared to the target object ($M = 29.28\%$, $SE = 1.86$; $t(19) = 3.10, p = .01, d = 0.99$) in the happy test trials only, signifying that they were able to encode the word-object associations only when the speakers were happy. During neutral test trials, attention to the target ($M = 32.09\%$, $SE = 2.07$) and non-target object was equal ($M = 32.07\%$, $SE = 2.23$; $t(19) = 0.004, p = 1.00, d = 0.001$).

Interestingly, when examining where infants had focused their attention during the happy and neutral familiarization trials, a repeated-measures ANOVA showed a significant main effect of AOI ($F(3,57) = 10.37, p < .001$; across emotions, mouth region > eye region; target > non-target) and a significant Emotion x AOI interaction ($F(3,57) = 3.29, p < .05$), which revealed that infants attended more to the mouth region in the happy trials ($M = 30.43\%$, $SE = 4.13$) compared to neutral trials ($M = 26.60\%$, $SE = 4.04$; $t(19) = 2.08, p = .05, d = 0.21$). Attention to the eye region and to both objects did not vary between emotions. Therefore, attending to the mouth may enhance infants’ word-object learning in a referential context. To that end, we also found a significant, positive correlation between infants’ MCDI productive vocabulary scores and attention to the mouth region ($r = .44, p = .05$), as well as a significant, positive correlation between infants’ MCDI receptive vocabulary scores and attention to the target object during neutral trials ($r = .45, p = .05$).

In conclusion, it is clear that happy speakers enhance infants’ word-object learning overall; however, possibly by increasing attention to the mouth area, infants can improve their word-object learning regardless of speaker emotion.
Individual differences in matched guise performance

The matched guise paradigm (Lambert et al. 1966) elicits subjective evaluation of a talker according to “guises” that differ minimally along a linguistic dimension. The primary use of matched guise experiments has been to interrogate potential social meanings of sociolinguistic variables. Campbell-Kibler (2008) shows that listeners can range widely in the meanings they assign to a variable, namely ING (workin’ ~ working). This study takes a known, widespread meaning for ING—that the /ɪŋ/ variant is associated with intelligence—and asks whether individuals differ in their sensitivity to this particular indexical association.

Following Campbell-Kibler, we manipulate naturally-occurring utterances for stimuli. 16 critical sentence-length utterances containing one instance of ING each, plus an additional 30 filler utterances not containing ING, were taken from unique talkers in the Philadelphia Neighborhood Corpus. Half of the critical utterances originally contained /ɪŋ/ and half originally contained /ɪn/, which we refer to as the “carrier style” conditions because other stylistic and referential elements in an utterance might co-occur with the original ING variant. The corresponding matched guises were created by splicing in the opposite variant from another instance of the same lexical item in the same talker’s interview.

Each participant heard both guises, 17-34 trials apart, for each critical utterance. The critical trials were counterbalanced for which guise was presented first and whether that guise was congruent with the original form of the utterance (the carrier style). The presentation order of the critical and filler items was pseudo-randomized within those constraints, then held constant across participants so that each individual received identical input. 34 participants rated each trial for “How smart does this person sound?” on a scale of 1 to 7. The resulting 1088 ratings were normalized within participant.

Mixed effects regression on the aggregate data indicates that the carrier style condition predicts smartness rating better than the presented guise, with no significant interaction. However, closer inspection reveals that pooling the data masks several distinct participant response patterns. 13 participants have the expected guise effect (/ɪŋ/ sounds smarter than /ɪn/) across utterances regardless of the carrier style, whereas 7 participants do have the expected guise effect but only in one carrier style or the other. A further 8 participants are not sensitive at all to either the guise or the original context. No participants are sensitive to the carrier style but insensitive to the presented guise.

These results suggest considerable, but not random, heterogeneity in how sensitive individuals are to sociolinguistic information. We identified three major response patterns: ING sensitivity, ING sensitivity only in certain contexts, and lack of sensitivity to ING. Such differences may inhere in individuals’ ability to perceive or attend to variation, propensity to adjust social evaluation based on linguistic behavior, interpretation of a particular variable’s indexicality, or a combination of these. The results also highlight the carrier phrase as integral to matched guise outcomes: the original variant in an utterance may be a reasonable proxy measure for the overall sociostylistic context in predicting social evaluation.


Attention and sociolinguistic salience

Most models of speech perception assume that attentional mechanisms are critical to speech processing and representation. In addition, recent work has shown that listeners encode social and linguistic information cued by phonetic variation in spoken words early and robustly. On this basis, we might assume that attentional mechanisms modulate encoding, resulting in the differential weighting of linguistic experiences, with social underpinnings. Here, we explore the role of attention in speech perception, with an eye toward socially-meaningful variation, within the dichotic listening paradigm.

Dichotic listening was introduced by Cherry (1953), who found that listeners can focus on one auditory stream while suppressing another, and that aspects of the suppressed stream may nevertheless attract attention. Listeners shadowed speech presented in one ear and ignored speech presented in the other. The ignored speech switched partway through, either from English to German, from forwards to backwards, or from a man’s to a woman’s voice. Listeners only reported explicit awareness of the switch when the sex of the unattended talker changed. This result has been interpreted to imply that, if the ignored stream captures attention, only its physical properties will be retained.

Two issues arise from this paradigm: (1) it ignores numerous social and linguistic differences between male and female voices; and (2) it uses gross measures of explicit recall, which may provoke null effects. Given current theory, the recent plentiful literature illuminating talker-specific effects, and the multi-tasking nature of being human, we might expect that listeners retain some degree of implicit memory for richly-detailed aspects of speech even in the absence of focused attention. Furthermore, early-activated social representations may modulate attention and the encoding of talker-specific episodic traces.

We extend Cherry’s paradigm to allow for fine-grained sensitivity to talker identity. We present a proof-of-concept pilot study establishing that detail of ignored speech is retained even when minimally attended and discuss the potential of this methodology to address issues of talker-specific attention and the weighting of linguistic events more broadly.

In our study, listeners performed a dichotic listening task. In the attended ear, they heard directions for a visually-presented map (Talker1). In the unattended ear, they heard 100 words which switched talker partway through (25 from Talker2, 50 from Talker3, 25 from Talker2). Critically, this task was immediately followed by an auditory lexical decision task featuring repeated post-switch words (Talker3) in random order. To establish whether small degrees of attention facilitate detailed encoding of speech, we tested whether listeners responded more quickly to early words near the “switch” than to late words farther downstream.

We ran two versions, each with 3 blocks ($N_1=27, N_2=22$). Pooling across blocks with mixed-effects regression, early words were recognized more quickly than late words ($p<0.05$). This result shows that listeners encode detailed information even under minimal attention.

Building on this result, we hypothesize that the degree of attention captured by the switch may vary by talker due to social and linguistic factors, affecting the extent of the recognition-facilitation effect reported here. More studies are underway to explore this hypothesis.
Can listeners recognize an emergent accent? Accent discrimination in the bilingual community of Galicia.

Minority language communities constitute complex linguistic settings where language choice is not always determined by the environment where speakers are raised, but can also be related to identity. For instance, in the bilingual community of Galicia, there is a group of speakers, neofalantes, whose dominant language is Spanish, but who decide to switch dominance to Galician at a late stage in their development (cf. O'Rourke and Ramallo, 2013, 2015). These so-called 'new speakers' have been documented in other European minority languages such as Irish, Welsh, Breton and Catalan (O'Rourke, Pujolar, Ramallo, 2015). However, the characteristics of these new speakers vary depending on the community, how they acquired the minority language, e.g., in a bilingual environment, through immersion educational programmes or as adult learners, and in terms of language exposure and use. Galician neofalantes are normally unbalanced bilinguals with higher proficiency in Spanish, but they switch to using Galician as their dominant language during adolescence for ideological or cultural reasons, sometimes displacing Spanish altogether.

Our previous work (reference supressed) investigated neofalantes’ production of Galician as compared to two control groups, Galician-dominant and Spanish-dominant speakers. Data from 56 participants was analysed (22 Galician-dominants, 20 Spanish-dominants and 14 neofalantes). Variables of interest were the Galician front /ɛ/-/e/ and back /ɔ/-/o/ mid-vowel contrasts, the reduction of unstressed word-final vowels, and the /s/-/ʃ/ contrast, which do not exist in Spanish. For mid vowels, an acoustic analysis showed that the neofalantes group produced intermediate categories that were different from those of Spanish and Galician-dominants, whereas their production of word-final vowels patterned with that of Galician-dominants. However, the neofalantes' fricative contrast was similar to that of Spanish-dominants. Overall, these findings suggest that neofalantes can acquire certain Galician features, but that there are limits to this flexibility.

One question that arises is whether these shifts in production are perceptible to listeners. The current study addresses this question using an accent identification task. Spanish-Galician bilinguals differing in their experience and use of Galician and Spanish (e.g., Spanish-dominant, Galician-dominant) listened to the same extract from a text produced by neofalantes, Spanish-dominant or Galician-dominant speakers, and indicated which group the speaker belonged to. They were subsequently asked to comment on whether particular factors had influenced their decision. The purpose of this experiment is twofold: first, to assess whether the production shifts made by neofalantes are sufficient for listeners in the speech community to identify the variety; second, to investigate whether accent identification varies as a function of listeners’ language background. One possibility is that listeners cannot differentiate between the neofalantes and Spanish-dominant varieties. Underlying category representations likely need to be tolerant to a range of acoustic variation (Evans & Iverson, 2007), and this might suggest that although neofalantes had made changes to their accent, these fell within this range and were thus unlikely to result in the emergence of a new neofalantes variety. On the contrary, accurate identification of the neofalantes’ accent would indicate the emergence of a new variety characterized by the effects of a switch in language dominance.
Listener knowledge about language variation: The case of variable (ING)

An extensive literature in sociolinguistics has demonstrated that English word final –ing (ING) realization (e.g. talking vs. talkin’) is conditioned by a host of linguistic factors, including the grammatical category of the (ING) word, its prosodic structure, and its phonological characteristics (e.g. Forrest 2015, Hazen 2008, Kendall 2013, Labov 2001, Tagliamonte 2004). Recently, research has shown that listeners are sensitive to variation in (ING), assigning different social judgments to speakers and utterances based on the use and frequency of (ING) forms (Campbell-Kibler 2007, Labov et al. 2011). More broadly, research on the processing of variation has further demonstrated that listeners interpret an ambiguous speech sample differently depending on the characteristics they attribute to the speaker (e.g. Niedzelski 1999, Strand, 1999). Thus, it is clear that indexical as well as linguistic information is useful to listeners in speech processing. However, the depth and detail of listeners’ knowledge about linguistic variation, and the relationship between listener knowledge and speaker knowledge (as listeners are, of course, also speakers), is still largely unknown.

There has been some previous work showing that listeners track characteristic pronunciations of phonetic features (e.g. VOT: Allen & Miller 2004, Clayards et al. 2008, Theodore 2014). However, monitoring such features serves an adaptive purpose in that it aids in comprehension; it is important to know whether a particular VOT value tends to correspond to /b/ or /d/ for a particular speaker. However, in the case of (ING), variability does not result in lexical ambiguity. In such a case, are listeners sensitive to variation in speakers’ use of (ING)? We hypothesize that, as conversational participants, listeners are motivated to monitor sociolinguistic variables for purposes beyond what is necessary for speech recognition (cf. Campbell-Kibler 2010, Labov et al. 2011, Levon & Buchstaller 2015). As a first step toward testing this hypothesis, we ask in this paper whether listeners can track such variation when explicitly asked to do so. We further ask whether this tracking is sensitive to the naturally occurring linguistic conditioning (i.e. grammatical and phonological effects) evident in production patterns—for instance, do listeners more accurately and quickly identify [ɪn] forms for progressive verbs than for adjectival forms, as would be predicted from production data?

To examine these questions, we conducted an experiment using a variant of a phoneme monitoring task, where participants pressed one of two buttons (–ing or –in’) each time they heard an (ING) word in a series of sentences produced by several distinct voices. Sentences were manipulated in a synthetic matched guise form (following Campbell-Kibler 2007) so that each participant heard one of two versions of each sentence, which were identical except for whether an (ING) word contained [ɪŋ] or [ɪn]. Analyses of participants’ accuracy and reaction times allow us to explore the extent to which listeners are sensitive to the conditioning linguistic factors of (ING). We conclude by considering implications for theories of speech processing and linguistic variation.


The role of duration in perception of vowel merger

Systematic differences in duration among vowel classes have been well documented, particularly between tense-lax vowel pairs (e.g., /u/ and /ʊ/) (e.g., Peterson & Lehiste 1960), though these durational differences have typically been considered phonetic rather than phonological in English. Recently, though, several studies have observed that durational differences among vowel classes may be more salient in both production and perception when vowel quality overlaps to a greater degree (e.g., Fridland, Kendall & Farrington 2014, Labov & Baranowski 2006, Ainsworth 1972), suggesting that duration may serve as a contrastive cue among qualitatively merged or near-merged vowel classes.

This study examines the effects of duration on perception of /u/, /ʊ/, and /o/ in pre-lateral contexts in two communities. Youngstown, a northeastern Ohio community, exhibits multiple patterns of merger and distinction among these phonemes, while Burlington, VT speakers are largely distinct. This study seeks to first examine whether duration can be utilized as a cue when discriminating between vowel classes. Next, this study investigates factors that potentially influence the degree to which duration is utilized in discrimination tasks, such as presence of merger/distinction in production/perception, exposure to merger/distinction in the speech community, and the ambiguity of the stimuli.

To determine which patterns of merger/distinction participants produced, target tokens were elicited in both connected and unconnected read speech from 33 native Youngstown speakers and a preliminary sample of 11 native Burlington speakers born 1933-2005. Merger/distinction was determined for each vowel class pair based on F1 and F2 measurements at 25% and 50% into the vowel-liquid sequence. To determine presence of perceptual merger, a multiple forced-choice discrimination task (target tokens=26, distractors=24) was administered in Praat.

Participants then took part in a second multiple forced-choice discrimination task, for which 26 target tokens from the Youngstown production data were used as the base stimuli to be manipulated. The vowel-liquid duration of each base stimulus was synthetically altered using the duration tier in Praat to produce four new stimuli, one for each duration category (99.5ms, 194ms, 288.5ms, and 383ms), yielding 104 total target stimuli; 46 distractor stimuli were also included. Stimuli consisted of both “ambiguous” and “non-ambiguous” tokens, determined by whether the speaker of each token produced statistically significant (i.e., $p < .05$) differences between vowel classes in F1 and/or F2.

Results suggest that participants consistently use duration as a cue not only for discriminating between the tense/lax vowel pair (/u/-/ʊ/) but even when discriminating between vowel pairs with additional featural distinctions (/o/-/ʊ/). Similarly, despite claims that duration may be utilized more extensively in cases of spectral overlap, there was no significant difference in the degree to which duration influenced answer choices for ambiguous and non-ambiguous stimuli. All participants utilized duration to some extent in the discrimination task, regardless of whether they exhibited merger in perception or production, or whether they are exposed to merger in their community. Durational cues appear to be independent of spectral overlap and accessible by both
merged and distinct speakers. Results provide implications for sociolinguists’ understanding of the production and perception of near mergers.

References
Listener and dialect effects in the false memory paradigm

In the false memory paradigm (Roediger & McDermott 1995), listeners are played a list of semantically related words where a centrally related word is absent (the "lure"), and are asked to recall what they heard. Researchers traditionally observe ~40% false recall of the unheard lure. Sumner & Kataoka (2013) report that the rate of such false memories is affected by the dialect that lists are presented in, finding that listeners have fewer false recalls with standard vs. non-standard dialects. In this paper, we investigate whether the effect of presentation dialect is dependent on various attributes of the listener.

73 native speakers of American English took part in this study. Participants were played 24 15-word lists taken from Stadtlter et al. (1999). The lists were recorded by six young males: two speakers of Standard American English (SAE), two speakers of Southern American English (STH), and two non-native, L1-Mandarin speakers (NNS). After a list played, participants were asked to write down all the words they could remember, and solved math problems between lists. Participants were scored on the number of words they correctly remembered, as well as the number of lures they incorrectly recalled. After completing the transcription task, participants answered a survey covering their demographics, their experiences as being (un)accented, their linguistic prescriptivism, their attitudes to Southern accents and Asian immigrants, their current mood, and the Big Five personality traits (Goldberg 1990). Factor analysis was used to reduce redundancy across questions.

The mean percentage of words correctly recalled was 42%. The lure was incorrectly recalled 40% of the time. We fit separate mixed effects models for the number of words correctly recalled, and the number of false recalls, testing the dialect of the speaker in interaction with various social attributes of the listener as independent factors. For the veridical recall model, listeners were significantly less accurate with NNS voices than SAE voices, and there was a main effect of the gender and order: men accurately recalled fewer words than women, and fewer words were accurately recalled as the experiment proceeded. In the false memory model, dialect only mattered in a marginal interaction with participants’ self reported mood: participants in a worse mood had more false memories, but only for the SAE voices. Order also mattered, such that there were fewer false memories as the task progressed. Note that these findings hold whether or not veridical recall is included as a predictor in the model (veridical recall is highly significant: the more words participants accurately recalled in a list, the fewer false memories they had).

In a post-task transcription test, participants were less accurate with the NNS and the STH speakers, suggesting that much of our results can be attributed to the intelligibility of the stimuli. The lack of primary differences between SAE and STH voices suggests that in this case the similar familiarity of both dialects for our participants was more critical than their different overt prestige. The effect of mood echoes findings from a listening in noise task using the same speakers (Walker 2015), where positive mood had a positive effect on performance, but only with Standard dialects.
Variation, social perception and linguistic alignment

When we talk, we dynamically adjust our linguistic behaviors (e.g., structural preferences, phonetic realizations) to become more (and sometimes less) similar to our interlocutors, a phenomenon known as linguistic alignment. Research in psycholinguistics has tried to understand linguistic alignment as the result of automatic processes, such as priming of linguistic representations (e.g., Pickering & Garrod, 2004) or implicit adaptation to the statistics of the environment (e.g., Bock & Griffin, 2000). Research in social psychology and sociolinguistics, on the other hand, has tried to understand alignment as a resource for navigating social interactions and achieving social goals (e.g., Balcetis & Dale, 2005; Pardo, 2006). In this presentation, we present results from an on-going series of experiments that investigate the joint influence of cognitive and social factors on alignment.

Recent work in our lab (Weatherholtz et al., 2014) provides evidence that people align their syntactic preferences with the talkers they encounter, even when they dislike these talkers or disagree with these talkers’ ideological beliefs. Specifically, participants were significantly more likely to use the prepositional object dative (PO: give [THE OBJECT] [TO THE RECIPIENT]) than the double object dative (DO; give [THE RECIPIENT] [THE OBJECT]) after hearing a talker produce PO structures in the context of a political diatribe about government spending (e.g., “Congress is giving [too much money] [to conservative special interest groups]”). The finding of overall syntactic alignment independent of social factors suggests that alignment is the result of automatic processes involved in language production. The strength of this overall alignment effect, however, was mediated by a range of social factors, including participants’ perceived similarity to the target talker (e.g., the extent to which they agreed with the talker’s political ideology) and participants’ individual tendency to compromise during conflict situations (e.g., when interacting with someone with different beliefs).

On-going work in our lab aims to understand the locus of these social effects. One hypothesis is that the linguistic processes involved in sequentializing preverbal thought (e.g., determining the relative ordering of the recipient and object phrases for dative sentences) interact with and are shaped by social cognition. Another possibility, however, is that the observed social effects on syntactic alignment reduce to attention differences across participants. For example, “compromisers” may attend more to information they disagree with than non-compromisers, and greater attention to a talker’s speech could explain greater alignment. That is, most psycholinguistic theories of alignment predict that greater attention to ambient linguistic information (e.g., a talker’s structural preferences) strengthens the encoding of this information in memory, which in turn makes this information more readily available, and hence more influential, during subsequent language production. In our current work in syntactic alignment, we investigate the extent to which social influences on alignment are due to a dynamic interaction between social and linguistic processes, or simply reduce to attention.

In addition to presenting the results of these on-going experiments, we discuss the statistical challenges that come with the high-dimensional data and large individual differences common in work on social perception. Further, we highlight recent advances in web-based experimentation that facilitate collection of spoken language and social perception data from large socially-heterogeneous samples.

References
Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning. JEP: General, 129.