Cognitive Mechanisms in the
Perception of Sociolinguistic Variation
By

BRANDON CONNER LOUDERMILK
B.A. (Reed College) 1994
M.A. (University of Texas, San Antonio) 2006

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Linguistics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

(David P. Corina), Chair

_____

(Robert J. Bayley)

_____

(Tamara Swaab)

Committee in Charge

2013

Abstract

In our increasingly multicultural and multilingual world, an understanding of how we perceive language, dialects, and linguistic variation and the relationship these features have to language attitude, plays an increasingly important role in shaping social behavior and policy. This study, situated at the intersection of cognitive neuroscience, psycholinguistics, and sociolinguistics, uses multiple methodologies to investigate how individual speaker/listener differences influence the perception of linguistic variation and the social attitudes it engenders.

The present study recorded participants' event-related brain potentials (ERPs) to investigate how and when information indexed by sociolinguistic variation is integrated during sentence processing. In a series of electrophysiological and behavioral experiments, participants listened to speech that varied in (ING) realization (working/workin'). In ERP experiment #1, Californian participants (n=28) heard short stories read by Californian speakers that were digitally manipulated to vary by speech register (formal/informal), variant (ING/IN'), and cloze probability (high/low). Findings showed (1) an N400 effect for cloze probability with low cloze words showing heightened negativities compared to high cloze words; (2) an N400-like effect for variant that was similar in latency, duration, and amplitude to the classic N400 cloze probability effect; with (3) increased N400 negativities for vernacular IN' variants compared to canonical ING variants; (4) increased N400 negativities to vernacular IN' variants preceded by informal register contexts compared to formal register contexts; and (5) greater N400 negativities for female compared to male listeners for sociolinguistic variant but not for cloze probability. According to these findings, word meaning and social information conveyed through spoken (ING) variants are integrated rapidly and concurrently into higher-order conceptual representations of meaning. However, when vernacular words variants are encountered during

language comprehension, these processes are taxed. Results from ERP experiment #1 provide evidence that the cognitive mechanisms that support language comprehension are sensitive not just to *what* is said, but also to *how* it is said and *who* hears it.

In experiment #2, Californian participants (n=19) listened to stories that were similarly manipulated to vary by register (formal/informal) and variant (ING / IN'). In order to further characterize the linguistic contexts under which sociolinguistic variants elicit N400-like negativities, passages were read by Californian and Southern speakers. We had reasoned that a stronger case for local linguistic contexts overriding broader sociolinguistic expectations could be made, if we could demonstrate an N400 like effect not just for vernacular IN' forms as in experiment #1, but also for canonical ING forms as well. Results showed (1) an N400-like negativity for IN' compared to ING words for Californian, but not Southern speakers; and (2) increased N400 negativities for formal ING passages uttered by Southern compared to Californian speakers. An additional analysis of the ERP data from experiment #2, investigated the role of listener gender. This analysis showed (1) N400-like negativities for females, but not males, listening to Informal IN' passages spoken by Californian speakers and Formal ING passages spoken by Southern speakers; and (2) heightened N400 negativities for females but not males listening to Californian speakers utter informal register contexts and for Southern speakers uttering formal register contexts

Taken together, the results of these experiments suggest that listeners have fine-grained representations of sociolinguistic and dialectal variation that are recruited during real-time language processing in order to help predict which variable word form a speaker is likely to utter. Importantly, these cognitive processes are further modulated by listener gender, with females showing particular sensitivity to sociolinguistic and dialectal norms. This work provides

compelling neural evidence that the cognitive mechanisms that support language comprehension are dependent on *what* is said, *how* it is said, *who* says it, and *who* hears it.

Acknowledgments

This work would not have been possible without the help, hard work, and expertise of numerous colleagues and friends. I owe a great deal of gratitude to my dissertation committee. My mentor Robert Bayley provided the original inspiration for this study and was invaluable for his breadth and depth of knowledge in variationist sociolinguistics. To Tamara Swaab, I give thanks for her expertise in and introducing me to the event-related potential technique. I am especially grateful to her for suggesting that I time lock my ERP data to the acoustic onset of the morpheme. I would also like to thank my adviser and dissertation chair David Corina for his psycholinguistic insights and his constant encouragement throughout this project.

I owe an enormous amount of thanks to the legions of research assistants who have assisted with the project over the year. First and foremost, I thank Richele Allen. This work *absolutely* would not have been possible without her unwavering enthusiasm and assistance with stimuli development and running subjects on various behavioral and electrophysiological experiments. I would also like to thank my other undergraduate research assistants, Christo Pallas, Victoria Brookins, Kyle Jones, Stephen Haight, Keith Penich, Tom Zolot, Max Stewart, and Timi Ihshii. I also offer my appreciation to fellow UC Davis graduate student, Serena Williams for her help with running ERP subjects over the summer months.

I would like to thank the numerous colleagues who have read and commented upon various aspects of this dissertation. I am particularly grateful to my good friend and colleague Vineeta Chand for "cracking the whip" and setting deadlines for this project. I also thank Lee Osterhout, William Labov, and Mike Grosvald for their insightful comments and criticism. I offer my gratitude to fellow sociolinguistic cognition blogger Christopher Stewart for reading

various drafts of this work. As well, I would like to offer my appreciation to Lauren Squires for reviewing my chapter on sociolinguistic cognition.

There are countless others at various conferences where this work has been presented that I owe my thanks too. In particular I would like to thank Greg Guy, Dennis Preston, Bridget Drinka, and Juliet Langman. Finally, I would like to thank the National Science Foundation for their generous support of Dissertation Improvement Grant 0952001.

Table of Contents

List of Tables

List of Figures

1. Introduction

This dissertation addresses a question that is germane to both psycholinguistics and sociolinguistics: What do listeners know about sociolinguistic and dialectal variation and how is this variation processed by the brain/mind? In a series of electrophysiological and behavioral experiments I address three sub-questions that comprise this overarching goal. 1.) What are the cognitive processes that support sociolinguistic variation perception during on-line language processing? 2.) Does information conveyed through speaker dialect influence variant processing? 3.) Do male and female listeners process variation differently during real-time language comprehension?

1.1 Perception of Sociolinguistic Variation

One of the central problems in the study of mind and language is how the listener recovers the intended message of a spoken utterance. From a cognitive science perspective this problem entails identifying: 1.) the stages of perceptual analysis between speech signal presentation and message recognition; 2.) the types of processing computations that occur at each stage; and 3.) the perceptual processing units and their representations in memory. Any work in psycholinguistics focusing on this issue, must address the question of how the listener's

mind/brain maps a raw acoustic signal into a meaningful message. Although researchers have been addressing this question for decades, relatively little progress has been made.

> At first glance, the solution to the problem of how we perceive speech seems deceptively simple. If one could identify stretches of the acoustic waveform that correspond to units of perception, then the path from sound to meaning would be clear. However, this correspondence or mapping has proven extremely difficult to find, even after some forty-five years of research on the problem. (Nygaard & Pisoni, 1995)

Many of the difficulties of speech perception result from what is known as the "lack of invariance" problem – there are few reliable and constant one-to-one mappings between the phonemes of a language and their phonetic implementation. Some of this variation is context dependent, for example, in English alveolar stops /t/ and /d/ are realized as flaps [ɾ] between vowels (e.g., the words <atom> and <Adam> are homophonous in most American English dialects). In addition, the speech sounds of a language are not discrete; rather they overlap with one another because of coarticulatory effects. For example, English /n/ which is normally pronounced with an alveolar place of articulation (POA), is realized with a dental POA when followed by the interdental fricative /ϴ/.

Finally, there exists both within-speaker and across-speaker variability. Different acoustic realizations uttered by the same speaker at different times can convey the same linguistic message; likewise different speakers produce different acoustic signals that can convey the same message (Staum, 2008). To cite just one example, in English the underlying phoneme /t/ can be

realized word as a fully articulated [t], a coarticulated glottalized form [ʔtˤ], and a singly articulated glottal stop [ʔ] (Sumner & Samuel, 2005). In addition to single speaker variability in language production, variation can also be observed across different speakers. For example, –er final words (e.g., 'baker'), are realized differently by the geographic region of the speaker. Speakers of the New York City dialect often produce r-less variants (bak[ə]), while speakers from other parts of the country produce r-ful variants (bak[ɚ]) (Sumner & Samuel, 2009).

Considering this massive variability in language production, the question that emerges, is how does the listener successfully map these different acoustic realizations and recover the intended message of the speaker? One factor that may aid language comprehension in the face of this variation, is that variability, though not categorically conditioned, does show probabilistic structure. Weinrich, Labov, and Herzog (1968) refer to this structure as "orderly heterogeneity" – that is, variation in phonetic realization demonstrates probabilistically regular patterns (Guy, 2007). Work in the field of variationist sociolinguistics over the past five decades has identified the systematic yet probabilistic constraints that condition variation in language production. These constraints include linguistically internal factors, for example, the lexical class or phonological environment in which a phone is realized. But they also include external social factors, such as the gender, race, socioeconomic status, and geographic region of the speaker in addition to the context and register of the speech act.

If listeners can store and access information about the probabilistic distributions of socially conditioned variation, then the process of real-time language comprehension would be greatly facilitated. Over the past decade, research in sociolinguistic cognition, has provided tantalizing evidence that suggests that listeners do indeed store and access information about

socially conditioned probabilistic variation. Early work by Niedzielski (1999), for example, showed that if listeners *believed* the speaker was an American or a Canadian, influenced their perception of vowel quality. Staum (2008) further demonstrated that perceived speaker race biased perception of ambiguous words induced by (-t/d) deletion (e.g., the form [mæs] can signify <mass> as in a church ceremony or it can signify <mast> as in a sailing boat, if the listener believes the speaker deleted the final [t]). Similarly, Koops, Gentry, and Pantos (2008) showed that the perceived age of the speaker influenced gaze fixation to words that are partially ambiguous for speakers with the PIN~PEN vowel merger (e.g., the first segments of the words <rinse> and <rent> are homophonous for speakers that merge /I/ and /ɛ/ before nasals). These and other studies have provided mounting evidence that listeners can use information about probabilistic sociolinguistic variation, posing difficulties for strictly bottom-up serial modular accounts of language processing (Fodor, 1983). The influence of non-linguistic information (e.g., speaker age, race, and origin) on these tasks, suggests that domain-general capacities and top-down contextual processes are involved in language processing under certain experimental manipulations. Although studies have shown that listeners can use sociolinguistic information in the service of certain off-line behavioral language tasks, relatively little is known about how sociolinguistic variation is handled in real-time during naturalistic language comprehension.

The present dissertation addresses this gap by investigating the perception of the single sociolinguistic variable (ING), which is realized in speech production with velar and alveolar nasal variants (e.g., *working/workin'*). Specifically, this dissertation presents a series of electrophysiological and behavioral experiments investigating how ING/IN' word forms are processed in real-time during language comprehension. Additional statistical analyses provide

further evidence for important gender/sex differences in processing the social aspects of language.

1.2 Organization of the Chapters

This dissertation is organized as follows. In Chapter 2, I introduce the sociolinguistic variable at the center of this study – variable (ING). In this chapter, I review the extensive variationist sociolinguistic literature on (ING). Here, I review studies that examine both the internal linguistic and the external social and stylistic constraints that are known to influence (ING) realization. The internal constraints that are discussed are the phonological environment and lexical class in which (ING) is realized. Next, I examine how (ING) pronunciation is influenced by register formality and social context. I also address the other social constraints on (ING) that have been identified, namely speaker gender, race, socio-economic status, and dialect. I additionally review studies examining language attitudes towards (ING). I conclude Chapter 2 by addressing how the various internal and external constraints identified in the literature informed my choices in experimental design and stimuli creation for the present project.

In Chapter 3, I move away from language production towards language perception. This chapter opens by introducing the newly emerging field of sociolinguistic cognition (Campbell-Kibler, 2010). Here I review recent work that has examined how sociolinguistic variation is perceived, processed, and represented. Specifically, I examine work showing how dialects are perceptually categorized and learned by naïve listeners and show how these processes are modulated by linguistic exposure. Following this I address how sociolinguistic stereotypes impact language processing. Specifically, I identify the extent to which listeners use extra-linguistic information (e.g., age, gender, and region) while processing variation in speech. This is

followed by a more general discussion of language attitudes toward sociolinguistic variation, and here I introduce one on the chronometric behavioral methods I employed in this research, the Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998). I conclude Chapter 3, by reviewing some recent and novel approaches to studying sociolinguistic cognition, including the event-related potential methods employed in this dissertation project.

Chapter 4 transitions from the world of sociolinguistics into the domain of psycholinguistics and cognitive neuroscience. In this chapter I examine two different psycholinguistic models of language processing that have been proposed, single and dual-step accounts of language perception and comprehension. After introducing the event-related potential (ERP) paradigm and the N400 ERP component, I assess these two different models of language processing by reviewing recent ERP studies that investigate how global contextual factors such as discourse, pragmatics, world knowledge, speaker identity, and extra-linguistic information, modulates the electrophysiological brain response. I conclude by arguing that the most parsimonious fit for these findings is a single-step model in which several levels of language representation are simultaneously processed during language interpretation.

In Chapter 5, I introduce the first of two event-related potential studies investigating the perception of (ING) during real-time language comprehension. In this first study, designed to assess the singe- and dual step models of language processing in regards to sociolinguistic variation, participants listened to spoken stories that were manipulated to create high expectations for specific words (cloze probability) and specific sociolinguistic (ING) variants. The questions that this experiment addresses are the following: (1) Do listeners have expectations for specific sociolinguistic variants as they do for word cloze probability? And (2)

Is social information conveyed by sociolinguistic variation processed within the same time window and in the same manner as word meaning itself?

In Chapter 6, I introduce a second event-related potential experiment designed to investigate how language processing is modulated by speaker dialect. Using a variation of my ERP paradigm, participants in this study listened to short passages read by speakers of different English dialects (i.e. Californian and Southern dialects). Specifically, this experiment tests whether listeners take into account the dialect of the speaker in order to predict which (ING) variant the speaker is likely utter. Although all speakers show (ING) variation in natural speech, the question addressed here is whether listeners make use of this probabilistic dialectal information to help facilitate on-line language processing?

In Chapter 7, I provide additional analyses of the ERP data from experiments #1 and #2, by examining the role of listener gender. Here, I investigate whether female and male listeners process (ING) variation differently. Specifically, I ask whether female listeners are more sensitive to the social dimensions of speech.

In Chapter 8, the final chapter, I summarize the results of these various experiments and analyses. In this final chapter, I propose several novel projects designed to address the limitations of the current study and pose several open questions for future research in the field of sociolinguistic cognition. Finally, I advance the conclusion that listeners have fine-grained mental representations of sociolinguistic variation that includes information about dialectal preferences. Moreover, I argue that listeners use these representations in the service of language comprehension by helping predict which variable form a speaker is likely to utter. I further argue that these cognitive processes are modulated by listener gender, with female listeners showing particular sensitivity to the social dimensions of language use. I conclude by proposing that

future psycholinguistic research need take into account not just *what* is said, but also *how* it is said, *who* says it, and *who* hears it.

2. The Sociolinguistic Variable: English (ING)

Since the pioneering work of Labov (1966), the subsequent four decades have witnessed an impressive accumulation of knowledge about language variation. Most of this research has focused on the patterning of sociolinguistic variables – structural units with different realizations (Chambers, 2009) or "a set of alternative ways of saying the same thing" that have social significance (Fasold, 1990); and which can have important social consequences such as discriminatory practices based on "linguistic profiling" (e.g., Baugh, 2000; Purnell, Idsardi, & Baugh, 1999). Research in this tradition has shown that linguistic variation displays "orderly heterogeneity" (Weinrich, Labov, & Herzog, 1968), in which alternating variants demonstrate probabilistically regular patterns (Guy, 2007). Numerous variationist studies have identified the systematic social, stylistic, and linguistic constraints that modulate the production of sociolinguistic variation. These include linguistically internal factors, such as lexical class and phonological environment, as well as external factors, such as gender, race, socioeconomic status, and register formality.

In defining the characteristics that make a "good" sociolinguistic variable, Labov (1972a) proposes several criteria: frequency, structure, and distribution. First, in order to uncover the systematic internal and external motivations in synchronic and diachronic variation, the variable

should occur frequently in natural conversation. Second, the variable should be structurally integrated "into a larger system of functioning units." Finally, the feature should be highly stratified, showing an asymmetric distribution across a range of independent variables (e.g., age, education, gender, etc.). In addition, a sociolinguistic variable recruited for perception studies, need also be easily measured and well established in the production literature in order to enable adequate hypothesis testing. For the current project, I chose to focus on a single variable that meets these criteria, English (ING).

2.1 Variable (ING)

One of the most well studied variables in sociolinguistics is English (ING) – the alternation of velar and apical nasals (e.g., *working* versus *workin'*) excluding monomorphemes (e.g., *ring*). Phonetically, the apical variant IN' can be realized as [In], [ən], or syllabified [n̩] (Trudgill, 1974), while ING is typically realized with a velar nasal [Iŋ] [1]. Most variationist production studies of (ING) collapse the various alveolar variants into a single IN' category allowing for a dichotomous examination of ING/IN' realization. Table 2.1 illustrates the basic pattern of ING/IN' alternation.

[1] Throughout this dissertation I use the following conventions: (ING) within parentheses refers to the sociolinguistic variable; ING without parenthesis refer to the canonical velar variant; and the vernacular alveolar variant is identified as IN' maintaining the typical (non-linguistic) orthographic eye dialect representation of "g-dropping."

Table 2.1 - Examples of ING/IN' alternation in spoken English. Reproduced from Tagliamonte (2012).

(a)	We were *having* [ŋ] a good time out in what we were *doin'* [n].
(b)	I've said there'd be *somethin'* [n] on her desk by nine o'clock tomorrow *morning* [ŋ].
(c)	I'm just *startin'* [n] at the *beginning* [ŋ] of this episode.
(d)	So they're *investigating* [ŋ] ways of *puttin'* [n] up a fire-wall.

Importantly, English (ING) is a widespread, yet stable sociolinguistic variable, which does not appear to be undergoing diachronic change. (ING) has been studied across a range of English speaking communities, both within the US (K. Campbell-Kibler, 2006; Hazen, 2008; Johnson, 2006; Kiesling, 1998; Labov, 1966, 2001b; Roberts, 1994; Shuy, Wolfram, & Riley, 1967; Wald & Shopen, 1985; Wolfram & Christian, 1976) and outside its borders (E. Douglas-Cowie, 1978; Reid, 1978; Shopen, 1978; Trudgill, 1974; Woods, 1978). Because (ING) has been so extensively documented from the perspective of production and language attitudes, it makes an excellent variable by which to investigate the perceptual processing of sociolinguistic variation.

In this chapter, I review the extensive literature on (ING). In section 2.2, I examine the internal linguistic constraints that are known to influence (ING) realization, specifically the phonological environment as well as lexical and grammatical constraints. Next, in section 2.3, I review studies that examine (ING) in relationship to style, register formality, and social context. In section 2.4, I briefly address the other social constraints on (ING) that have been identified in the literature, namely gender, race, socio-economic status, and geographic region. Next, in

section 2.5, I review studies that have examined language attitudes towards (ING). In section 2.6, I briefly address how the various internal and external constraints on (ING) informed the experimental design and creation of stimuli. This chapter concludes with section 2.7 which briefly summarizes (ING) in relationship to the present research.

2.2 Internal Linguistics Constraints

Over the past decades, a number of variationist studies have identified the internal linguistic constraints of (ING). These include phonological, lexical, and grammatical constraints that influence the likelihood that (ING) will be realized in either its canonical ING or its vernacular IN' form.

2.2.1 Phonological Constraints

As noted by Campbell-Kibler (2006), there are relatively few phonological influences on (ING) realization. First, the variable is only realized in unstressed syllables, as stressed syllables have near categorical ING realization in most language varieties. In terms of non-categorical phonological constraints, both regressive assimilation and progressive dissimilation have been noted. Following velar and alveolar stops favoring ING and IN' respectively, and preceding velar and alveolar stops favor IN' and ING (Cofer, 1972; Houston, 1985; Roberts, 1994). Labov (2001a), however, has argued that there is no strong phonological conditioning of (ING) for following velars or apicals.

2.2.2 Lexical Constraints

In addition to phonological constraints, specific lexical effects have also been observed, namely certain *thing* words (e.g., something, nothing, everything, anything) have been shown to influence rates of (ING) realization. The words *something* and *nothing* favor IN' realization, whereas *everything* and *anything* near-categorically favor ING. It has been argued that these lexical effects may be an actual reflection of stress effects, as the words *anything* and *everything* has secondary stress on (ING) compared to unstressed *something* and *nothing* (Cofer, 1972; Houston, 1985).

2.2.3 Grammatical Constraints

A number of studies have shown that that grammatical category influences rates of (ING) realization. The basic finding of these of studies is that verbal forms favor IN' and nominal categories favor ING, with gerunds falling between these extremes. Houston (1991), for example, examined nearly 20 grammatical categories of (ING) realization, identifying six significant categories: proper nouns, derived nominals, monomorphemic nouns, gerunds, progressives, and prepositions. As further collapsing of categories did yield significant results, Houston argues that grammatical conditioning of (ING) is best understood along a continuum rather than as discrete categories. Similarly, work by Labov (2001b) on Philadelphia speech communities has provided additional insight into the grammatical conditioning of (ING). His studies identified five statistically significant grammatical categories: progressive, participle, complement, gerund, and

gerundive nominal. Labov argues that the clustering of verbal and nominal categories best supports grammatical conditioning as a discrete binary function rather than as a continuum of categories (Fig. 2.1).

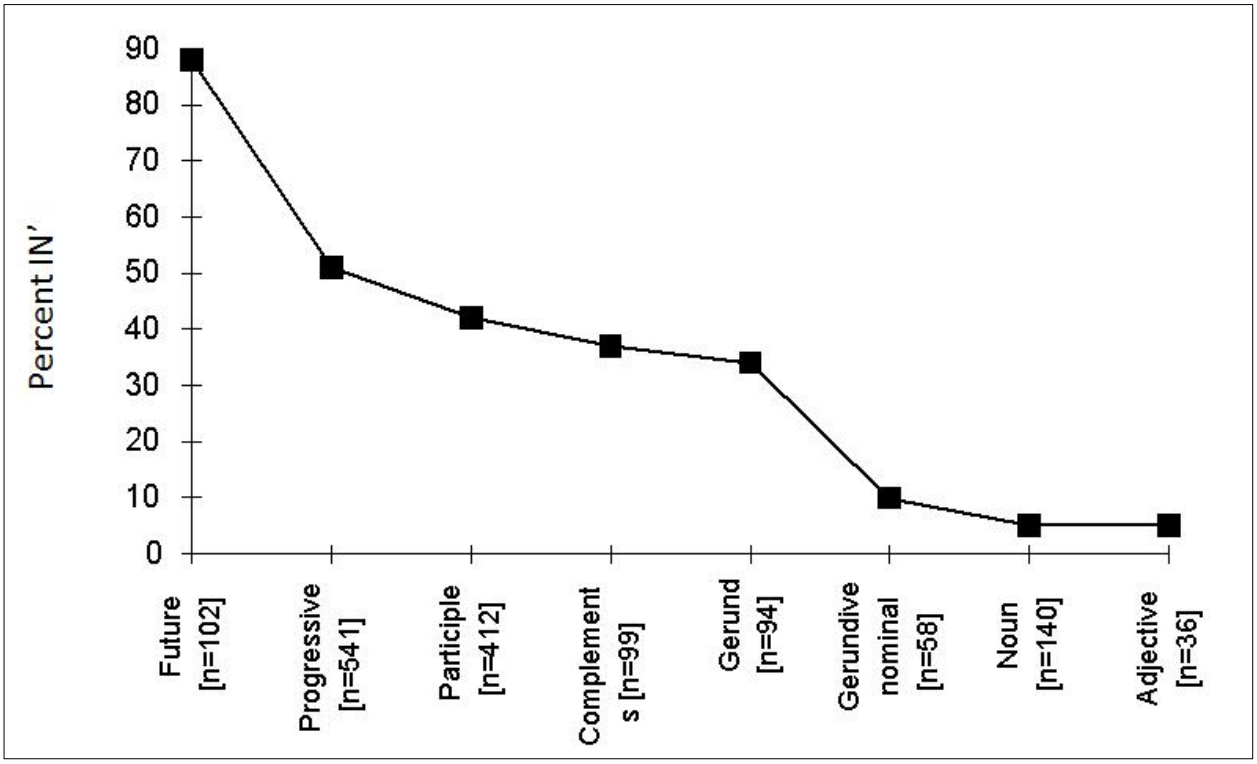

Figure 2.1 – Percent of IN' realization by grammatical category of the stem for 33 speakers in lower middle class neighborhoods in Philadelphia (Labov, 2001).

Both Labov (1989) and Houston (1985) have examined the grammatical conditioning of (ING) from a historical perspective. The basic story is that in the transition from Old English to Modern English, several invariant morphological forms merged into a single contemporary suffix, and that the present day grammatical conditioning of (ING) reflects the historical grammatical differences of these original morphemes. In Old English there were three distinct suffixes: a masculine derivational suffix *–ing*, a feminine derivational suffix *–ung*, and the

present participle inflectional suffix *–ind* (Houston, 1985). By the end of the 13th century, the feminine derivational suffix *–ung* had fallen out of use, with the *–ing* derivational suffix being preferred for both male and female forms. Similarly, by the 15th century the *–ing* suffix began to replace the original *–ind* inflectional suffix. Houston argues that this historical alternation between *–ing* and *–ind* suffixes partially accounts for the contemporary grammatical conditioning of Modern English (ING). This basic argument is that the specialization of alveolar and velar variants began to be lost in Early Middle English, but that because the merger was incomplete, competition remained between these forms, which can still be observed in the contemporary preference of IN' for verbal forms and ING for nominal forms. As noted by Campbell-Kibler (2006), since the nineteenth century (ING) has been used with essentially the same variants and social associations. She speculates that its consistent internal constraints and social associations are due to its historical stability as a case of "stigmatization without change" (Labov, 1966).

2.3 Stylistic Constraints

One of the strongest constraints on (ING) is speech style – the relative formality of the speech context, which according to Labov correlates with the amount of "attention paid to speech" by the speaker. Although studies define "formality" differently, the basic observation is that speakers change the way they speak and the variants they use in different social contexts. Consider, for example, a conversation between friends compared to a lecture delivered to an audience. In the lecture, a more formal environment, we expect to encounter more canonical variants as speakers are more conscious of their speech and their audience. In contrast, in a

casual conversation between friends, we expect the interlocutors to use relatively more vernacular variants compared to the more formal environment.

A number of studies have reported strong correlations between (ING) and speech register or context. In early work, for example, Fischer (1958) examined the speech of a ten year old boy across three different contexts: a formal language task, a formal interview, and an informal interview. He reported almost categorical use of ING in the formal language protocol which shifted across tasks until a prevalence of the vernacular IN' variant was observed in the casual interview. Labov (1966), using four different "styles" – casual speech, formal speech, reading, and word list, reports (ING) stratification similar to that observed by Fischer (1958), with increasing rates of IN' with decreasing "attention paid to speech" (Fig. 2.2).

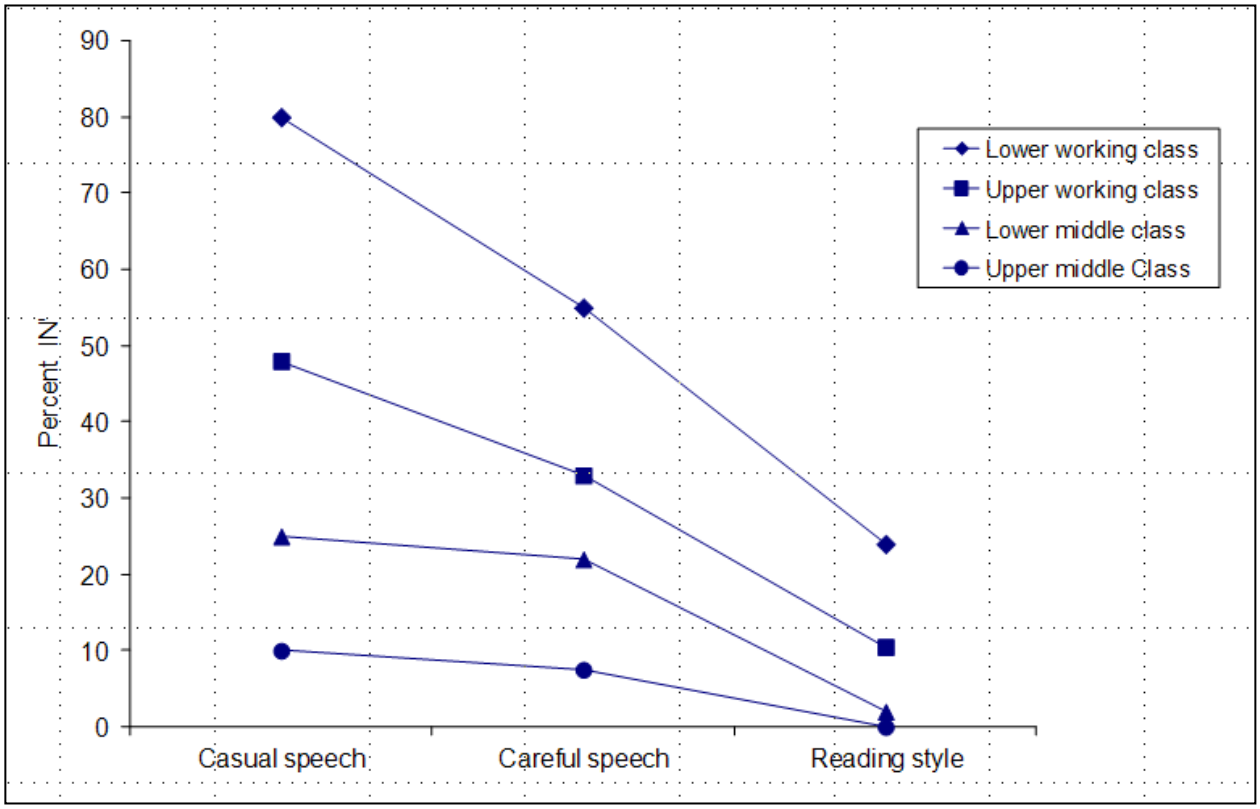

Figure 2.2 - Social and stylistic stratification of (ING) in lower east side New York City adults. Reproduced from Labov (1966).

Similarly, Trudgil (1974), in his study of English in Norwich, England, reported increasing IN' variants as participants moved from more formal styles to "everyday" speech. Kiesling (1998), in a study of men's fraternity reports similar stylistic constraints on (ING). Reid (1978), with adolescent informants in Edinburgh, also notes increasing IN' use as the conversational topic became less focused on language and adult material (Table 2.2).

Table 2.2 - Percentage of IN' in Edinburgh youth, by speech situation. Reproduced from Reid (1978).

Situation	**% IN'**
Reading Passage	14
Interview	45
Peer Session	54
Playground	59

Douglas-Cowie (1978), in a study on Northern Ireland, used interlocutor familiarity as a proxy for speech formality. In the less formal context participants were interviewed by the investigator, a local resident familiar with the interviewees. In the more formal context, participants were interviewed by an unfamiliar interviewer who spoke with an RP accent. Douglas-Cowie reported increasing IN' variants when speaking to a familiar interviewer compared to the unfamiliar RP interviewer. Cofer (1972) operationalized formality by

manipulating interview topic and timing (with the assumption that participants are more nervous during the initial part of an interview compared to latter parts). The first part of the interview as formal by timing and by topic (e.g., biographical information and questions about language use); the middle part was informal by timing and topic (e.g., childhood games, great danger); and the final section was informal by timing, but formal by topic (e.g., describe how to tie a shoelace, define a successful man). Results showed relatively high ING use in the initial formal section and higher IN' use with the second section, but in the third section (ING) realization interacted with speaker socioeconomic status and race, with white middle-class speakers returning to high ING use, but working class speakers using relatively greater IN' variants. Finally, Wald and Shopen (1985), relying on natural speech (rather than speech elicited through interviews or tasks), defined formality in terms of interlocutor familiarity, reporting greater IN' use when speaking with friends compared to family members and others (Table 2.3).

Table 2.3 - Percentage of IN' by conversational context. Reproduced from Wald and Shopen (1985).

Speaker	Joking Style	Arguing Style
Male	46	24
Female	28	21

Taken together, these studies show consistent stylistic constraints on (ING), namely increased use of the vernacular IN' variant in less formal speech contexts whether defined by interview topic, timing, task, social relationship, or some combination of these factors.

2.4 Social Constraints

In addition to linguistic and stylistic constraints, (ING) realization is also influenced by social correlates. Although (ING) has been studied in a range of social and geographical contexts, the most thoroughly documented constraints are socio-economic status and gender. Accordingly, in the next two sections, I detail research showing how (ING) is influenced by these factors, before turning to examine some of the less-well documented social correlates (e.g., race, region, and age).

2.4.1 Socioeconomic Status and Education

Of researchers who have examined (ING) in relationship to the socioeconomic status (SES) and social class, most have reported a strong positive correlation (e.g., Cofer, 1972; Labov, 1966). Although researchers have operationalized SES and class differently, the basic finding is that lower SES speakers tend to use the IN' form, while higher SES speakers use the canonical ing variant. The basic pattern is illustrated in Shuy et al. (1967) and Reid (1978) which are reproduced in Fig. 2.3 and Fig. 2.4 respectively.

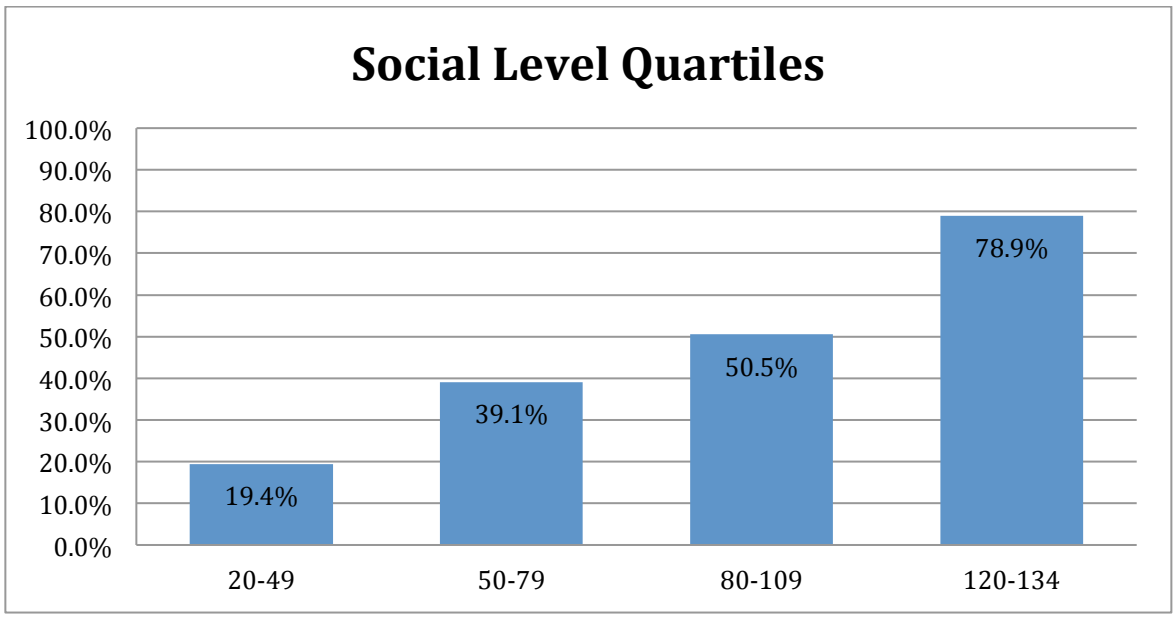

Figure 2.3 – Percent of IN' by social level (Detroit). Reproduced from Shuy et al. (1967).

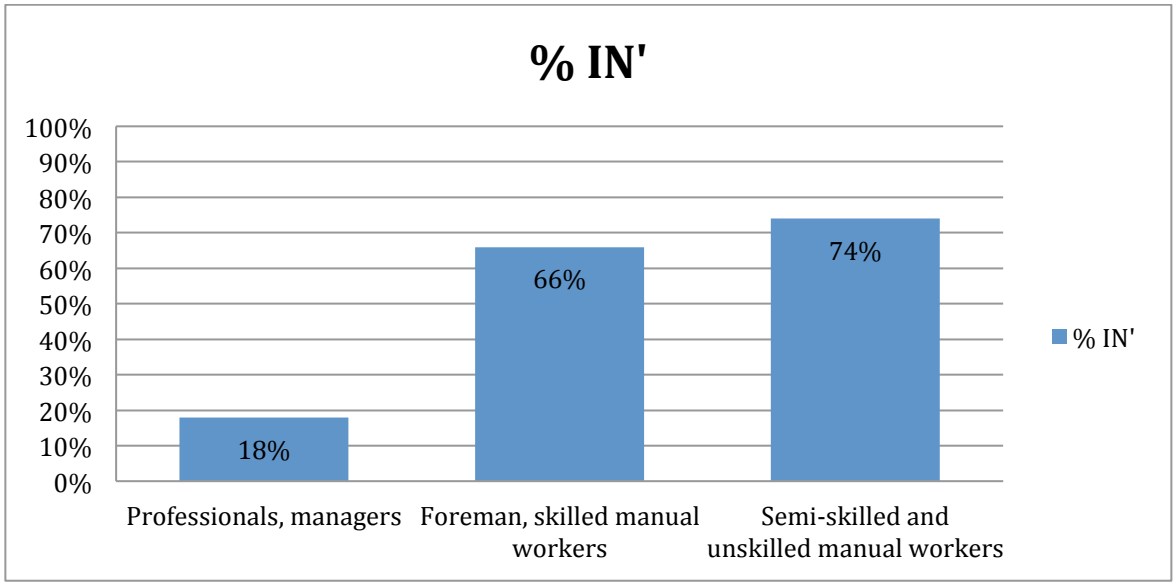

Figure 2.4 - Percentage of IN' (Edinburgh) by father's occupation. Reproduced from Reid (1978).

Among his Philadelphia informants, Labov reports (ING) is influenced by both speaker education and occupation in careful speech, but occupation only in casual speech (Labov, 2001b). In

summary, most studies report an inverse relationship between SES and IN' realization, such that as social status increases, IN' use decreases.

2.4.2 Gender

A number of variationist production studies have observed that (ING) realization is constrained by speaker gender. In an early study, Fisher (1958), who studied child-rearing practices in a New England village, noted statistically significant differences by gender, with females using the ING variant at higher frequencies than the males. A study by Shuy, et al. (1967) on Detroit speech noted a similar effect of speaker gender, with females using ING more frequently than males. Similarly, Trudgill's study in Norwich, England, noted greater IN' use among males compared to females. Wald and Shopen's (1985) study in Canberra, Australia also reported greater IN' realization for male speakers compared to females. They also noted effects of listener gender on (ING) realization, observing that females increased IN' usage when addressed by male interlocutors, suggesting that women "have a greater awareness of the stylistic effect of the style shifting" (p.535). In general, these studies show a preference for canonical ING variants for women speakers compared to men. Table 2.4 shows a summary of gender and (ING) production.

Table 2.4 - Percentage of IN' pronunciations by speaker gender across studies and locale. Adopted from Campbell-Kibler (2006).

Study	**Location**	**Speaker Gender % IN' tokens**	
		Male	**Female**

Shuy et al. (1967)	Detroit	62 %	21 %
Shopen (1978)	Canberra	24 %	16 %
Houston (1985), <35 yrs	U.K.	88 %	72 %
Houston (1985), >35 yrs	U.K.	78 %	76 %
Wald and Shopen (1985)	Canberra	23 %	24 %

2.4.3 Race & Ethnicity

A small number of studies have examined (ING) realization and race, typically focusing on white and black speakers, with a common observation that black speakers tend to use more IN' variants compared to white speakers. For example, Shuy et al., (1967) reported about 20% greater IN' use among black compared to white speakers. As illustrated in Table 2.5, Labov's (1966) New York study also reported greater IN' use among black compared to white speakers. Importantly, these effects were mediated by whether the speaker was a New Yorker or an out-of-towner, with black out-of-towners showing the most IN' use, but no significant differences among white speakers for residency status. Levine and Crockett (1966), in a study conducted in North Carolina, reported race-based (ING) differences that were larger than educational differences. Although speaker race often interacts with other social factors, most studies show increased use of the vernacular IN' variant among black compared to white speakers.

Table 2.5 - Percent of IN' pronunciations by speaker race, residency status, and language task (i.e. interview vs. reading). Reproduced from Labov (1966).

Speaker Race & Residency Status	Language Task % IN' tokens	
	Interview	**Reading**
White New Yorkers	31 %	13 %
White out-of-towners	37 %	8 %
Black New Yorkers	62 %	18 %
Black out-of-towners	77 %	42 %

2.4.5 Region

A number of studies have examined the relationship between (ING) realization and speaker region. Labov (1966) reported greater vernacular IN' use among Black out-of-towners compared to Black native New Yorkers. Wald and Shopen, analyzing natural conversational data, observed greater IN' use among residents of Los Angeles (24.8%) compared to Canberra (19.8%). Others have noted increased vernacular IN' use in the Southern United States. Holding speaker education constant, Mencken (1963), reported greater IN' use among educated Southerners compared to Northerners or Easterners. Wolfram and Christian (1976) report higher rates of alveolarization among Appalachian speakers compared to other speakers, including those from the non-Appalachian south. These observations have led Labov (2001b) to argue that in the Southern United States, "the /in/ form is used almost exclusively in speech, even in the most formal kind" (p.90).

2.5 Language attitudes about (ING)

Although most (ING) studies in sociolinguistics have focused on language production, a number of researchers have focused on language attitudes and ideologies. From a psycholinguistic perspective, these are important because different social values and attitudes held by a community, as well as individual mood, have been shown to modulate language processing (e.g., De Goede, et al., 2009; Van Berkum, Holleman, Murre, Nieuwland, & Otten, 2007). In terms of language attitude, Wald and Shopen (1985), analyzing interview and self-report data, classified participant responses to ING/IN' into three categories: personally negative (e.g., uneducated, lazy), socially lower (e.g., working class, blue collar), and neutral descriptions (e.g., everyone says it sometimes). The researchers found that participants consider ING to be the standard or canonical form, and that people who used the IN' variant were described as personally negatively and socially lower. In order to elicit language attitudes, Campbell-Kibler (2007) used a matched-guise task in which she manipulated naturally spoken passages that differed in (ING) realization. She found that the vernacular IN' variant increased the perceived strength of Southern accents (e.g., lack of education, "redneck") and that the ING variant increased the perceived "gayness" of the speaker (e.g., lowered masculinity, metrosexual).

In recent work, Labov and colleagues (2011) used a matched-guise task that manipulated (ING) variants in passages read by a "newscaster-in-training." In a series of experiments that manipulated the frequency of (ING) variants in read passages, subjects were asked to rate the speech on measures of professionalism. Across these experiments, the researchers found similar overall results, namely a logarithmic progression between the frequency of vernacular IN' and the negative evaluation of speaker professionalism (Fig. 2.5). Consistent with the production

literature (Labov, 1966; Trudgill, 1974), they also reported an effect for gender in which females evaluated the marked variant more harshly than males. Moreover, subjects from South Carolina were found to be more tolerant of IN' use than subjects from Philadelphia. In an experiment manipulating speaker dialect, they found no differences in overall evaluation, suggesting that the frequency of (ING) alone was the determining factor in eliciting these overt attitudes. Taken together, these studies show support for Labov's claim that (ING) is a sociolinguistic stereotype and that the vernacular IN' variant is generally associated with negative attitudes.

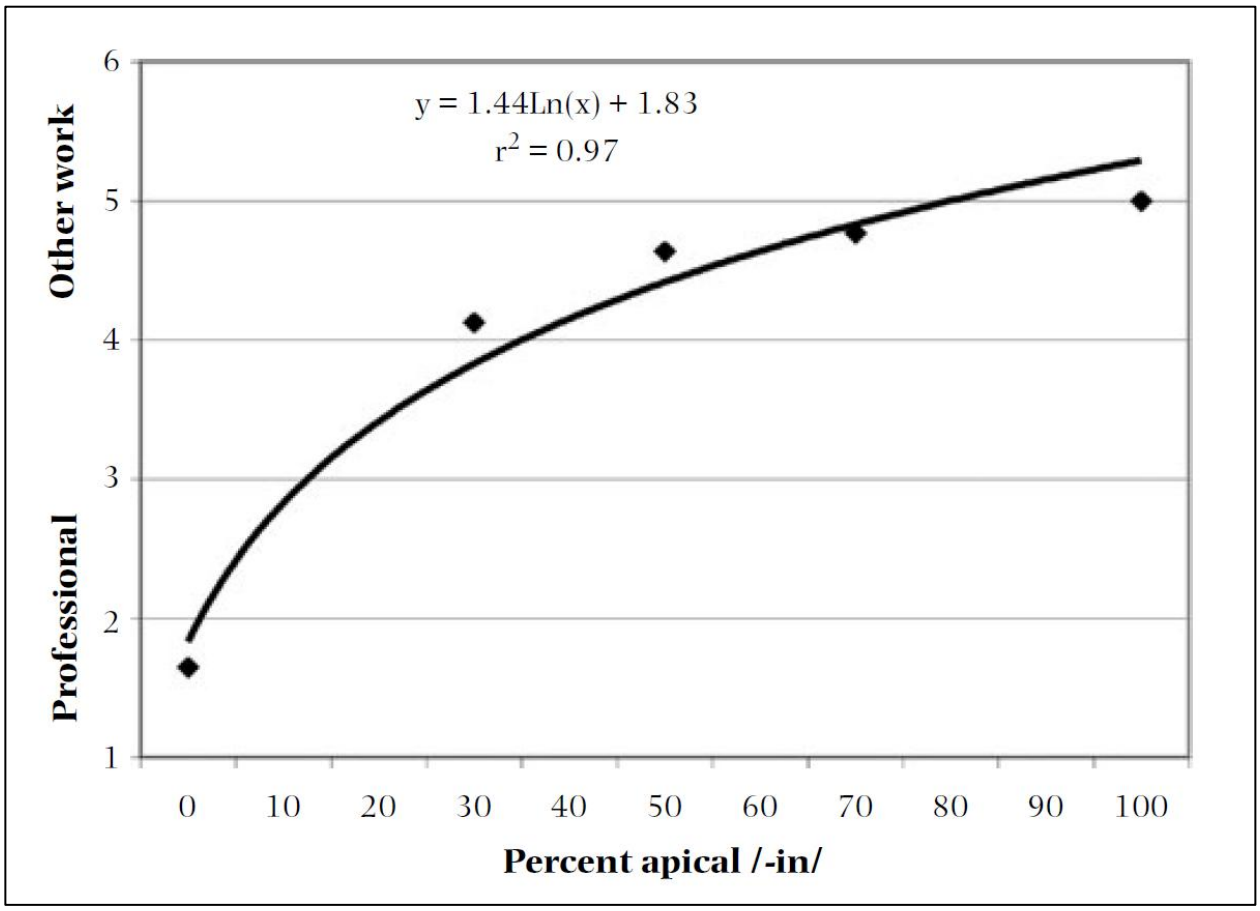

Figure 2.5 - Percentage of vernacular IN' variants in news passages and perceptions of professionalism. Adopted from Labov et al., (2011).

Importantly, recent work has been conducted on attitudes towards (ING) using methods adapted from social psychology, namely the Implicit Association Task (IAT) developed by Greenwald and colleagues (Greenwald & Banaji, 1995; Greenwald, Banaji, & Nosek, 2003; Greenwald, McGhee, & Schwartz, 1998). Briefly, the IAT is an experimental chronometric technique developed to measure the strength of association between two dichotomies. Traditionally, the IAT has been used to study a range of social biases including racial and gender stereotyping (Amodio & Devine, 2006; Dasgupta, McGhee, Greenwald, & Banaji, 1999). More recently it has been adopted to explore issues of sociolinguistic interest (e.g., Babel, 2010; Pantos, 2010). During IAT experiments, subjects sort stimuli representing four concepts into two response categories (Greenwald, et al., 2002), and the difference in mean response latencies across the two key mappings (the D value) is argued to measure the strength of the association and its automatic evaluation. Recently, Campbell-Kibler (2012) used this technique to measure individuals implicit associations between (ING) and its typical social and linguistic correlates. In her first experiment, she tested participants using written (ING) forms (e.g., *making/makin'*) contrasted with northern/southern states, white/blue collar professions, and news anchors/country singers. Results showed statistically significant D values in the predicted direction (Table 2.6)

Table 2.6 - IAT Results for written (ING) vs. States, Professions, and Singer/Anchors. Reproduced from Campbell-Kibler (2012).

		D mean	*p*-Value
(ING) vs.	States	0.38	<0.001
	Professions	0.44	<0.001
	Singers/anchors	0.24	0.002

In her second experiment, Campbell-Kibler used audio (ING) tokens in lieu of the written forms used in her first experiment, and contrasted them with northern/southern states, /ay/ monophthongization, and /t/ release. Results showed statistically significant associations in the predicted direction between (ING) and States and /ay/ monophthongization, but not with /t/ release (Table 2.7).

Table 2.7 - IAT Results for audio (ING) vs. States, /ay/, and /t/. Reproduced from Campbell-Kibler (2012).

		D mean	*p*-Value
(ING) vs.	States	0.30	0.002
	/ay/	0.18	0.004
	/t/	0.09	NS

In summary, results from Campbell-Kibler (2012) show that the IAT is sensitive enough to capture associations between (ING) and geographic region and other social and linguistic correlates previously identified in the literature. Additional details on the IAT and how it was adapted to the current project can be found in Chapters 3 and 9.

2.6 Studying (ING) from a psycholinguistic perspective

As with most sociolinguistic variables, production of (ING) is constrained by a number of internal linguistic and external stylistic and social factors. In the present section, I briefly address the rationale behind the experimental design in relationship to these numerous sociolinguistic

factors. In conducting psycholinguistic perceptual experiments, controlling these myriad influences becomes prohibitively expensive in terms of experimental design and stimuli creation. This is especially the case for event-related potential studies which require substantial trials per condition using unique stimuli (i.e. words) for each trial. This is even more important when studying the N400 component because studies have previously shown an attenuation of component amplitude for repeated words. Although methodological details are provided in subsequent chapters, here I briefly describe the general approach I used for the two event-related potential experiments reported in this dissertation.

In choosing unique (ING) words for these experiments, I only chose verbal forms (e.g., progressives) which typically show the most variability in ING/IN' realization (approximately 50% as illustrated in Fig. 2.1). I refrained from using nominal forms because they are near-categorically realized as ING, and I sought to strike a balance between ecological validity and experimental control for investigating sociolinguistic cognition. Importantly, this is a very conservative approach, and I expect that using nominal forms would have provided stronger results, perhaps supporting the contention that sociolinguistic variants used in disfavored environments carry more social weight (Bender, 2001), and which accordingly may tax cognitive processes to a greater extent. Similarly, for sentence-final critical words, I refrained from using the lexical item *going*, which shows near categorical realization as *goin'* or *gonna*. I did, however, use *going/goin'* variants in non-critical positions in order to establish realistic discourse contexts. Likewise, I avoided the use of *thing* words (e.g., *something*, *nothing*) in critical positions.

In terms of syllabic stress I only choose (ING) words in unstressed or secondarily stressed environments, avoiding stressed environments which are categorically or near-categorically

realized with the velar ING variant in most language varieties. Rather than control the preceding phonological environment, which would have been near impossible using (ING) words with extremely high and low cloze probabilities, I chose instead to randomly assign trials and their associated (ING) words to experimental groups. Because of this random assignment, the previous phonological environment should not bias any observed results

In addressing social constraints I similarly sought to balance ecological validity and experimental control. All speakers were roughly balanced for age, socio-economic status, and age. In experiment #1 (Chapter 5), I equally balanced male and female speakers, but in experiment #2 (Chapter 6), I used only male speakers from two separate dialect regions. Again, this was a very conservative approach, as previous researched has shown that female speakers, compared to males tend to avoid the use of vernacular IN' variants.

2.7 Summary of (ING)

The sociolinguistic variable (ING) is one of the most well-studied variables in the variationist paradigm, having received extensive attention from researchers for nearly five decades. English (ING) is a sociolinguistic "stereotype" – a form of variation of which speakers are aware, and which is open to overt commentary (e.g., Southerners drop their G's) and stigmatization (Labov, 2001b). Research on (ING) has identified a number of linguistic, stylistic, and social constraints that influence whether the variable is likely to be realized in its canonical ING form or in its vernacular IN' form. For purposes of the current project the most important of these constraints are register formality, gender, and geographic region. In consideration of these constraints,

informal register contexts, males, and Southern dialects all promote vernacular IN' realization compared to formal contexts, females, and speakers of more "standard" English varieties.

Variable (ING) provides an excellent choice for examining aspects of sociolinguistic cognition and language processing from a psycholinguistic perspective. Although word variants such as *working* and *workin'* "mean" the same thing, they carry different social weight and are associated with different social correlates. This is particularly interesting in terms of language processing and representation. Are these different word forms represented separately in the mental lexicon? Are language processes, such as lexical access and sentence-level integration differentially impacted by these variable word forms? Does the language processing system consider one word form to be more prototypical or canonical than the other? While processing these variants, does the language system take into the broader sociolinguistic context, such as the dialect of the speaker? Do individual differences in variable production impact variant perception? Does an individual's strength of language attitude towards (ING) facilitate or hinder language processing? Are female listeners more sensitive to the social aspects of language use? These are just a few of the many questions that emerge while studying sociolinguistic cognition and language processing.

3. Studies in Sociolinguistic Cognition

Foreword

The present chapter reviews recent experimental work in sociolinguistic cognition. It was originally published as "Chapter 7: Psycholinguistic Approaches" in the Oxford Handbook of Sociolinguistics (B. Loudermilk, 2013). Although the text of this chapter remains unchanged from the original, it has been reformatted with different section numbers and page numbers for consistency with this dissertation. For citation purposes, please cite the original Oxford Handbook chapter.

The factors influencing sociolinguistic variation have been well described across languages and dialects, yet open questions remain: How is sociolinguistic variation perceived during real time language processing? What is the nature of the cognitive representation? How do linguistic and social stereotypes influence perception? A more nuanced understanding of these mechanisms would play an important role in addressing some of the fundamental issues in sociolinguistics, including how children acquire social constraints; how interlocutors 'converge' on similar styles; and how individuals develop attitudes towards speakers of different language varieties. This chapter seeks to highlight some of the key developments and methods used in the emerging field of *sociolinguistic cognition* (Campbell-Kibler, 2010).

3.1 Sociolinguistic Cognition

The fundamental goal of the study of sociolinguistic cognition is to characterize the computational stages and cognitive representations underlying the perception and production of sociolinguistic variation. Work in this field has focused on three issues: (1) understanding the manner and degree to which social knowledge, beliefs, and stereotypes influences perceptual processing of spoken language; (2) identifying the cognitive stages, representations, and time course of processing; and (3) characterizing the way in which language and language variants engender attitudes and bias language processing.

This chapter is organized into four sections. Section 3.2 discusses different methods for examining how dialectal variation is represented, perceived, and learned. Section 3.3 reviews studies investigating the role of sociolinguistic stereotypes in speech processing. Section 3.4 explores the attitudinal aspects of language variation by presenting two recent studies using innovative variations of the matched-guise technique. This section concludes by introducing a promising technique, the implicit association test, which may be able to address some of the limitations of alternative methods. Section 3.5 reports on studies that use eye-tracking and event-related brain potentials to investigate sociolinguistic cognition.

3.2 Dialectal Variation

One important line of research involves characterizing what naïve language users know about dialectal variation and how they come to learn it. Two approaches are used: (1) perceptual categorization studies in which participants listen to recordings of speakers of different dialects and attempt to categorize them; and (2) perceptual learning experiments which attempt to

characterize how dialectal variation affects learning.

3.2.1 Perceptual categorization of dialects

In perceptual categorization experiments participants listen to speech that contains dialect-specific features (e.g., vowel variants) and then categorize these variants according to perceived speaker origin. In order to tease out the effects of prior dialect exposure on perception, participants with different regional backgrounds are recruited (Clopper, 2010).

Forced-choice Perceptual Categorization

During forced-choice perceptual categorization tasks, participants listen to speech and then identify "where they think the speaker is from" by selecting from a predefined dialect category. Preston (1993) used this technique to investigate how regional background influences dialect classification. Participants from Michigan and Indiana listened to speakers from nine different cities in the United States on a north-to-south continuum. Although participants had difficulty distinguishing between north and midland speakers, they were better at distinguishing between northern and southern speakers. Preston also reported an effect for participant region: Michigan listeners distinguished northern from southern speakers at a more northern boundary than did the Indiana listeners. Early results such as these underscored the importance of linguistic exposure in cognitive processes.

More recently, Clopper and Pisoni (2004b) examined how geographical mobility influenced dialect identification. Here, participants listened to speech from six dialect regions and then classified that speech into one of six dialect regions presented on a map. Participants comprised two groups: "homebodies" who had only lived in Indiana, and "army brats" who had

lived in three or more states. On average, participants performed quite poorly on this task (~30% accuracy), but were still statistically above chance performance. Army brats outperformed homebodies, and army brats who had lived in a given region categorized talkers from that region more accurately than army brats who had not lived there.

While listeners seem to perform rather poorly on forced-choice dialect classification tasks, these findings highlight the importance of linguistic background by showing how exposure to regional variation affects how well regional accents of unfamiliar talkers are identified. One of the primary limitations of the forced-choice task is that the number of sorting categories and category labels are predefined by the researcher. As such, there may be a poor fit between to participants' actual mental representations of dialectal variation and the categories they choose from.

Auditory free classification

One technique that overcomes a limitation of forced-choice classification is the auditory free classification task (C.G. Clopper, 2008). In this task, participants are presented a screen that allows them to play speech snippets. The graphical representation of the sound stimuli can then be dragged on a grid to form participant-defined groupings of speakers who "speak the same dialect" (Fig. 3.1). Aggregate results can be interpreted by submitting them to clustering analysis (which identifies how objects group together perceptually) and multidimensional scaling (which identifies the most salient perceptual dimensions). Importantly, the number of categories and category labels need not be pre-specified, thus eliminating one limitation of forced-choice categorization.

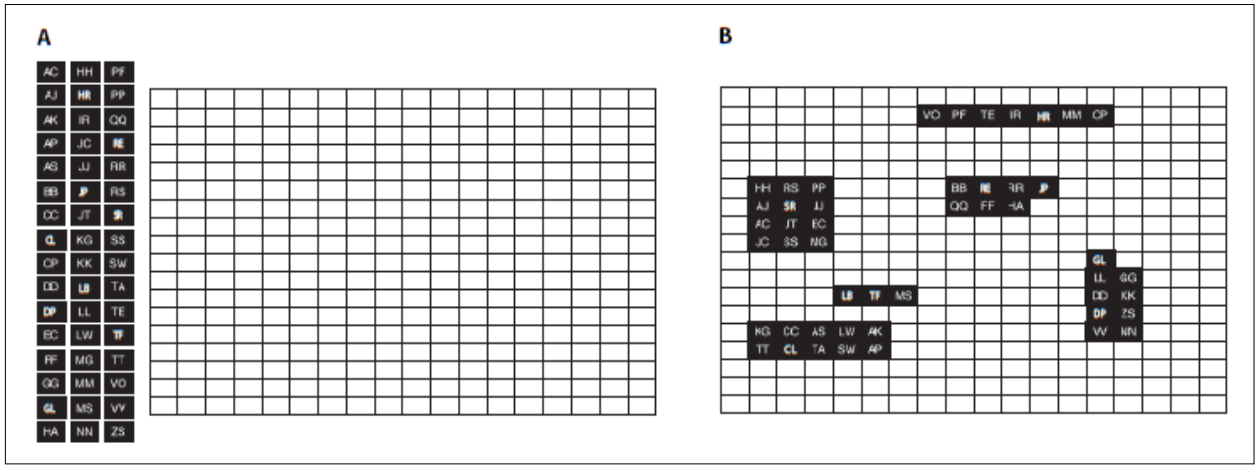

Figure 3.1 - Free classification task, with 48 stimulus items. A 16 x 16 grid (A) before the task and (B) after the task (Clopper, 2008).

Given these more relaxed constraints, Clopper and Pisoni (2007) examined how listeners would freely classify the same speakers from the six dialect regions previously used in the forced-choice experiment (C.G. Clopper & Pisoni, 2004b). Results showed that participants classified speakers into an average of ten dialect groups, suggesting that listeners do indeed have the ability to make finer discriminations of regional varieties. However, despite this, clustering analysis identified three broad dialect categories (i.e. New England, South, and Midwest/West) which corresponded to the three categories identified by their previous forced-choice study. Multidimensional scaling revealed that linguistic markedness, geographic region, and gender were the primary dimensions by which listeners classified dialects.

Synthesized speech

Rather than use natural speech, some researchers use synthesized speech in order to exercise greater control over their experimental manipulations. Plichta and Preston (2005) administered a perception test to subjects from 20 regions of the United States. Their study focused on a

southern dialect feature, monophthongization of /ay/, and sought to measure participants' ability to use acoustic information to distinguish speaker origin on a north-south continuum. Seven resynthesized vowel tokens were embedded in a [gVd] frame, to create words ranging from a fully diphthongal 'guide' to fully monophthongal 'god.' Participants listened to these words and clicked a map to indicate one of seven locations on a north-south continuum. Although participants reported being unable to discriminate between tokens, results showed otherwise. There were significant differences in "southernness" between every step of the continuum, with the most monophthongal tokens being rated most southern and the fully diphthongal tokens being rated most northern.

Plichta and Rakerd (2010) investigated how listeners' region of origin influences the perception of a change in progress, the Northern Cities Chain Shift (NCCS). In the NCCS, multiple vowels are shifted: /a/ fronting, /æ/ fronting and raising, and /ɛ/ lowering. Participants were recruited from two different Michigan cities: Ishpeming, a working-class town in the Upper Peninsula (UP), and Detroit, a metropolitan city in Lower Michigan (LM). Importantly, residents of Detroit exhibit advanced stages of the NCCS whereas residents of the Upper Peninsula do not. The researchers set to determine (1) whether vowel tokens subject to NCCS /a/ fronting would be interpreted differently depending on whether the tokens were preceded by speech from either a LM speaker or a UP speaker and (2) whether listeners' linguistic background impacted perception.

By varying the frequency of F2, the researchers synthesized a continuum of seven vowel tokens ranging from /a/ to /æ/, which were embedded in /hVt/ and /sVt/ syllable frames to create words that varied from 'hot' to 'hat' and 'sock' to 'sack'. Participants listened to sentences containing the ambiguous words spoken by either an LM talker or a UP talker (e.g., 'Bob was

positive that he heard his wife, Shannon, say: [hVt]'), and classified the ambiguous word as either 'hot' or 'hat.' Upper Peninsula residents, who have limited exposure to NCCS speech, showed no statistical difference in crossover points for speaker (Figure 3.2, left). In contrast, Lower Michigan participants highly familiar with NCCS speech, showed a significant speaker effect. When Lower Michigan listeners heard the UP speaker, they discriminated at a crossover point similar to the one used by Upper Peninsula listeners for both UP and LM speakers (Figure 7-2, upper right). However, when Lower Michigan listeners heard speech from an LM talker, they shifted their discrimination threshold a full step further to the front (i.e., a higher F2 value) (Figure 2, bottom right). These results demonstrate that information about a speaker's dialect plays an important role in vowel identification which is influenced by dialect exposure.

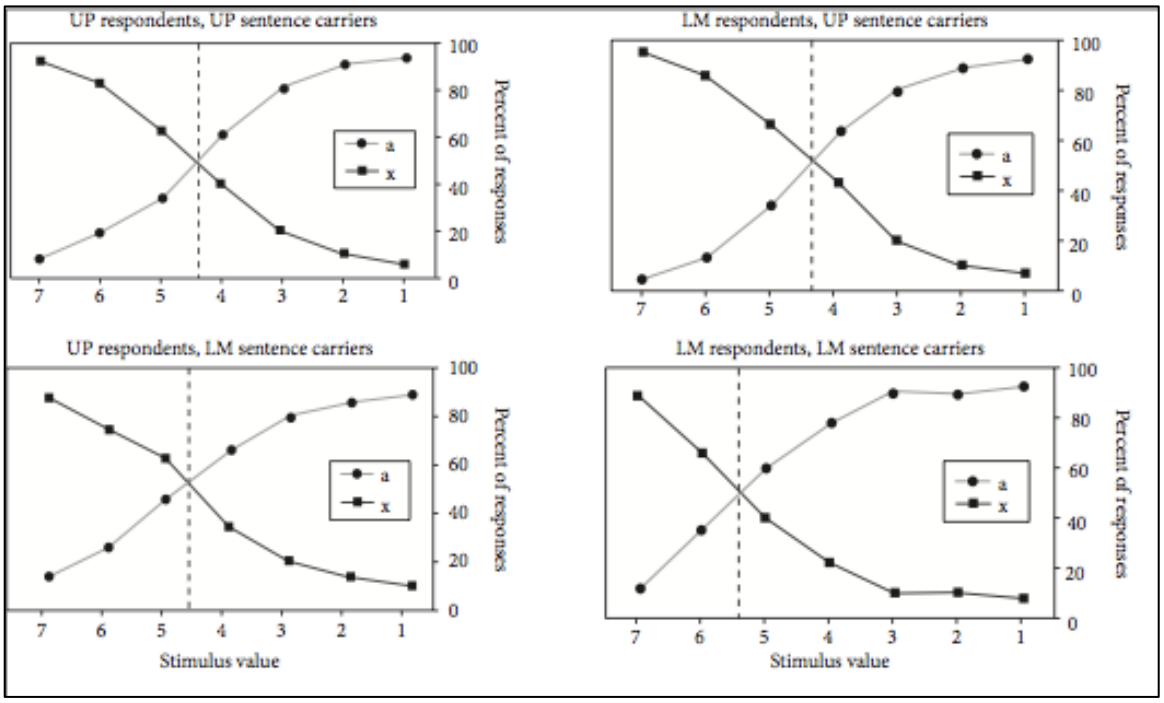

Figure 3.2 - Upper Peninsula listeners (left) show no difference in perceptual boundaries for ambiguous /a/-/æ/ tokens. Lower Michigan listeners (upper right) use the same perceptual boundary /a/-/æ/ as Upper Peninsula speakers when they encounter non-NCCS speech. Reproduced from Plichta and Rakerd (2010).

Priming

Sumner and Samuel (2005, 2009) investigated how dialectal variants are processed and represented over different time spans. In their 2005 study, they investigated how the realization of coda /t/ affects language processing. In the Long Island dialect of American English, word-final /t/ is variably realized as a fully articulated [t], a coarticulated glottalized form [ʔtˤ], and a singly articulated glottal stop [ʔ]. In order to determine the immediate and long-term processing consequences of these variants, the researchers used an auditory semantic priming paradigm. In this paradigm, participants listen to prime-target word pairs and make a lexical decision on the target word. The prime word 'flute' for example primes, or speeds the reaction time, for the *related* target word 'music' compared to the *unrelated* target word 'money.' In this study, participants heard prime words consisting of the three regular variants of a word (e.g., basic [flut], coarticulated [fluʔtˤ], and glottal [fluʔ]). Also included were primes with an arbitrary feature mismatch (e.g., [flus]) that paralleled the place of articulation mismatch for the glottal variant (e.g., [fluʔ]). Target words were semantically related or unrelated to the prime (e.g., 'music' or 'money'). Results showed that participants were equally primed for all three variants compared to the feature mismatch condition. Interestingly, this shows that at least in terms of immediate processing, there are no cognitive costs associated with processing non-canonical regular variants.

A second experiment by Sumner and Samuel (2005) used a long-term repetition priming paradigm, in which all prime words were presented in a first block of trials and all target words were presented in a second block of trials. This experiment additionally differed from the first, in

that it used a repetition priming paradigm in which target words were repeats of prime words. Prior work has shown that participants respond faster and more accurately if they have already encountered the same stimulus. Using this design allowed the researchers to investigate the degree to which one variant primes another variant relative to itself. The temporal lag in the blocked design allowed them to identify which form is encoded in memory over the intervening interval. Results showed that of all the possible combinations of variants, only the Basic[t] – Basic [t] prime-target pairs showed a strong priming effect. Although the partially glottalized variant is more frequent in actual production, it did not show a priming effect. This evidence argues against frequency-based exemplar models of representation, suggesting that the canonical form serves as the underlying representation.

In a subsequent study Sumner and Samuel (2009) investigated the effects of dialectal experience on the perception and representation of phonetic variation, specifically the alternative realization of –er final words (e.g., 'baker'). Speakers of the New York City dialect (NYC) often produce r-less variants (bak[ə]), while speakers from other parts of the country (General American - GA) produce r-ful variants (bak[ɚ]). Three groups of participants were identified: (1) Overt-NYC participants were third generation New York City residents who were primarily r-less; (2) Covert-NYC participants were second generation New York City residents who were r-ful ; and (3) GA participants who were born and raised outside New York City and were r-ful in their speech.

To examine how well dialectal surface features are processed by these groups, several priming experiments were conducted. As expected, results showed that the r-less variant posed difficulties for the GA speakers but not for the Overt-NYC participants. Interestingly, the Covert-NYC speakers patterned with the Overt-NYC speakers on the immediate perceptual tasks

(Fig. 3.3, left panel). Specifically, both groups showed facilitated priming effects for the r-less variant in the short-term repetition priming and the semantic priming experiments. However, in contrast to the Overt-NYC participants, the Covert-NYC speakers showed no r-less priming effect on the long-term repetition priming experiment, in which primes presented in blocks were preceded by targets about 20-30 minutes beforehand (Fig. 3.3, right panel). Apparently, for these Covert-NYC speakers, only the GA form is encoded in memory; the Overt-NYC speakers appear to encode both variants of final –er equally well. These findings suggest there are multiple ways of "having" a dialect: in speech production, in speech perception, and in long-term encoding.

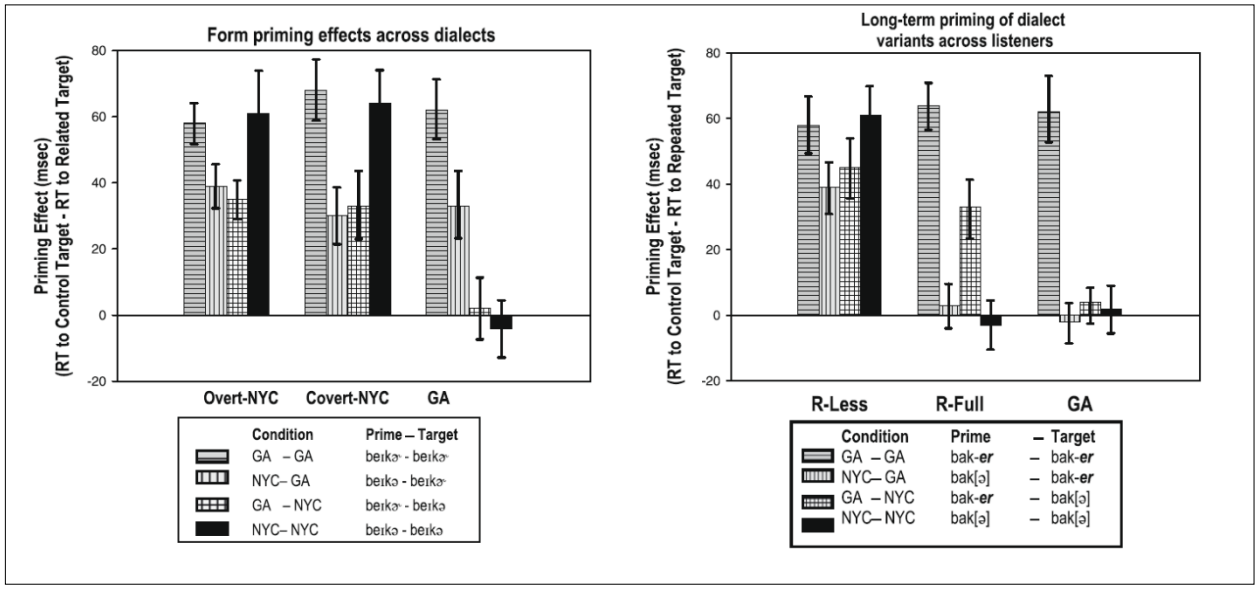

Figure 3.3 - Short term repetition priming by dialect group (left panel). Long-term repetition priming by dialect group (right panel). Reproduced from Sumner and Samuel (2009).

Squires (2011) used a syntactic priming paradigm to explore how social information such as talker gender and social status affects sentence processing. Previous syntactic priming studies have shown that priming participants with one syntactic alternant (e.g., passive voice) makes that

alternant more likely in production and increases expectations during perception. Squires' experiments tested whether exposing participants to a nonstandard agreement pattern (e.g., *The dog don't bark; There's dogs in the yard*) in the prime sentence made it more likely that participants would perceive nonstandard agreement in the subsequent target sentence. Prime trials included the full realization of the subject-verb construction with a non-subject noun masked. Critically, target trials masked the subject noun, and participants chose between singular and plural noun photos (e.g., *dog/dogs*).

Squires found that priming was effective for these sociolinguistically variable agreement patterns: exposing participants to nonstandard agreement in a prime stimulus increased the likelihood that they interpreted the target stimulus as nonstandard. Because syntactic priming is argued to reflect speakers' cognitive representations of linguistic structures, these results suggest that speakers store knowledge of grammatical variation as part of linguistic competence and apply that knowledge during processing. Squires' experiments also found differential priming effects of the gender, social status, and individual identity of the prime and target talkers, suggesting that at least some social knowledge is also recruited during sentence processing.

3.2.2 Perceptual Learning

Recent work has investigated how exposure to phonetic variation can influence “perceptual learning” and alter the nature of linguistic and sociolinguistic representations. Perceptual learning is indicated by improved performance following language exposure, and the studies reported here directly bear on issues germane to sociolinguistics, speech perception, and dialect learning. Perceptual learning experiments generally comprise two parts: an initial learning phase followed by a testing phase. During learning, subjects perform a lexical decision task while they

hear an assortment of control words, non-words, and critical words that contain an ambiguous segment.

A study by Maye, Aslin, and Tanenhaus (2008) examined how listeners adjust perceptual boundaries after brief exposure to a novel dialect. Participants first listened to a 20-minute recording of the "Wizard of Oz" read in a synthesized standard American English accent, after which they were given a lexical decision test. Several days later, subjects were presented the same 20-minute story, but this time read in a synthesized dialect in which all the front vowels were systematically lowered (e.g., *witch* → *wetch* [wɛʧ]). Again, after exposure, participants were given a lexical decision test. The authors found a nearly 20% increase in acceptance rate for *wetch* items after exposure to the novel dialect compared to the pre-exposure test. This increase was significant even for stimuli that did not occur in the story, suggesting that adaptation is not purely lexical in nature. The findings highlight the rapidity with which listeners may re-map their perceptual spaces in order to accommodate the speech they encounter.

Clopper and Pisoni (2004a) also used a perceptual learning dialect categorization task to investigate the role of speaker variability on the learning of six regional varieties of English. Two groups of participants were used: one-talker listeners and three-talker listeners. During training the one-talker participants listened to sentences read by a single talker representative of each of the six dialects. In contrast, the three-talker participants were exposed to the speech of three talkers representing each of the six regional varieties. The three talkers from each variety were chosen by the researchers to be good, medium, and poor representatives of their regional dialect. During training, subjects indicated dialect region on an interactive map, with feedback highlighting the correct response. Training was divided into three blocks: (a) talkers reading the same sentence multiple times, (b) talkers reading an alternative multiple times, and (c) talkers

reading a number of different novel sentences.

Following training, the two groups of participants were tested on their ability to (1) identify the training talkers on novel sentences, and (2) generalize their dialect knowledge to novel speakers. The single-talker listeners outperformed the three-talker listeners on all training phases and the test phase. However, the three-talker group outperformed the single-talker group on generalization to novel speakers. This cross-over effect suggests that talker variability encountered during perceptual learning makes it more difficult to learn specific exemplars, but that it facilitates more robust dialect representations that enable learners to generalize these features to novel speakers.

3.3 Sociolinguistic Stereotypes and Language Processing

Over the past decade, a number of studies have examined how sociolinguistic stereotypes impact language processing. These studies seek to identify the extent to which listeners use extra-linguistic information (e.g., age, gender, and region) while processing speech. The questions addressed by this research are: what sort of tacit knowledge do naïve listeners have regarding variation and its social and linguistic correlates; and how does this knowledge influence language processing. These studies manipulate the perceived social characteristics of the speaker, while holding the auditory stimulus constant. This allows researchers to tease apart which aspects of social information listeners rely upon while attending to speech.

Perceived Speaker Origin

In an early study, Niedzielski (1999) examined the relationship between regional linguistic stereotypes and speech perception by using a vowel matching task. Her study focused on Canadian Raising, a phonetic process in which the /aw/ diphthong (as in the word *house*) is pronounced with the tongue raised and fronted compared to canonical /a/. Canadian Raising is a linguistic stereotype that the residents of Detroit hold about the speech of Canadians. Canadian Raising is also found in the speech of White, middle class Detroit residents, but they themselves do not report any awareness of this in their own speech.

In this study, Detroit area participants heard sentences and were asked to choose from a list of re-synthesized vowels, the token that best matched the vowel in a critical word. All of the sentences were spoken by the same Detroit speaker who had raised /aw/ in her speech. However, whereas one group of subjects believed that the speaker was from Detroit, the other group believed that the speaker was from a nearby Canadian city. When subjects believed a Canadian had spoken the word *house*, 60% of the participants chose the raised token. In contrast, when participants believed a fellow Detroiter had spoken the same word, only 11% chose the raised /aw/ token. The results for all /aw/ words were similar: if the speaker was perceived to be Canadian 53% chose the raised token in contrast to the 15% who chose the raised token for the perceived Detroiter. This suggests that perceived speaker origin is sufficient to bias phonological processing.

In a similar study by Hay and colleagues (2006), participants listened to the speech of a New Zealander but were provided an answer sheet with the word 'Australia' or 'New Zealand' written on the top of the form. The study focused on the /I/ vowel which is realized differently in the two dialects –raised and fronted in Australian English and centralized in New Zealand

English. Findings were similar to those reported in Niedzielski – when given the Australia answer sheet, listeners perceived more raised and fronted variants than when given the New Zealand answer form. Interestingly, these effects were observed even when listeners reported that they knew that the speaker was in fact a New Zealander. This suggests that instead of a dialect expectation bias per se, these results might be more attributable to a low-level priming mechanism in which "mere exposure" to the concept of Australia is enough to induce a perceptual shift.

Perceived Speaker Race

A study by Staum Casasanto (2008) examined whether listeners use knowledge of speaker race to resolve lexical ambiguities. Her study focused on coronal stop deletion, a characteristic of the speech of African American English (Rickford, 1999). Stimuli were sentences that contained a lexical ambiguity caused by deleting an underlying /t,d/. For example, the word 'mast' pronounced with a deleted –t and the word 'mass' are homophonous and thus ambiguous without a supporting context. Participants were visually presented pictures of black or white faces while they listened to short passages containing one of these ambiguous critical words. In one trial they heard

1.) The [mæs] probably lasted...

The audio clip then was followed by a visually presented continuation, either:

a. through the storm

b. an hour on Sunday

The "through the storm" continuation is more likely with a deleted-t interpretation of [mæs], while the "an hour on Sunday" continuation is more consistent with no deletion. Participants

responded by button press to indicate whether the continuation was plausible.

Findings revealed a significant interaction of speaker race and deletion. Participants responded faster for continuations that were consistent with a deleted-t (the *mast* interpretation) when accompanied with a black face. Likewise, participants responded faster for a non-underlying –t continuation (the *mass* interpretation) when accompanied with a white face. These findings demonstrate that listeners use knowledge of correlations or probabilities related to race and variation.

Gender Stereotypes

Work by Strand and Johnson (1996) has examined how perceptual processes are influenced by gender-based stereotypes. Their first study examined the discrimination boundary between /ʃ/ and /s/, the latter of which has a higher frequency spectrum reflective of the smaller size resonating cavity in front of the constriction. Importantly, variation in spectra for men and women is often attributed to relative differences in vocal tract length. Four speakers whose voices were judged to be prototypically male/female and non-prototypically male/female were recorded. Ambiguous /s/-/ʃ/ tokens were generated and spliced to a following VC coda spoken by each of the four speakers to form the words *sod* and *shod*. Participants listened to these ambiguous stimuli and categorized them.

Results showed that the discrimination boundary was determined by both the gender and the gender prototypicality of the speaker, with the largest differences observed between prototypical male and prototypical female voices. A second experiment accompanied the ambiguous tokens with video clips of a male or female face. Results showed that the same fricative was perceived differently depending upon whether it was accompanied by a male or

female face. Results showed that the /s/-/ʃ/ boundary was shifted up for a female face and shifted down for a male face. These results show that complex social expectations can influence relatively low-level cognitive mechanisms, challenging purely bottom-up modular accounts of processing.

3.4 Language Attitude

In the study of language attitude researchers have typically relied on three broad approaches: the societal treatment approach, which typically involves participant observation and ethnography; the direct approach in which participants are directly questioned about their language beliefs and attitudes; and the indirect approach, or matched-guise technique (Garrett, Coupland, & Williams, 2003; Lambert, Hodgson, Gardner, & Fillenbaum, 1960). The societal treatment approach is qualitative in nature and has been criticized on the grounds that the researcher is inferring attitude from behavior. Likewise, the direct approach has been criticized on "whether subjects' [overt] verbal statements of their attitudes" accurately reflect their "underlying disposition" (Knops & van Hout, 1988). The matched-guise technique (MGT) developed by Lambert and colleagues (1960) addresses these short-comings by trying to covertly elicit language attitude.

Matched Guise Experiments

In MGT studies, participants listen to repeated passages of speech read by a single speaker under different guises (e.g., language, dialect, etc.), and are asked to make judgments on some aspect of the speaker (e.g., intelligence, friendliness, etc.). Because the speaker and the content of the recording are held constant, and because the listener *believes* they are listening to different

speakers, any attitudinal differences can be attributed to the manipulation (i.e. guise) under study.

Although the MGT has been used to investigate attitudes towards gross features of language, finer distinctions can be made. Labov et al. (2011), for example, conducted a series of experiments that manipulated the frequency of (ING) variants – the alternation of velar [Iŋ] and apical [In] pronunciations (e.g., *working* vs. *workin'*). Subjects were asked to rate the speech of a newscaster-in-training on measures of professionalism. Across excerpts, the ratio of *–ing* and *–in* tokens were manipulated. The researchers found a logarithmic progression between the frequency of apical *-in* and the negative evaluation of speaker professionalism. Consistent with the production literature (Labov, 1966; Trudgill, 1974), they also reported an effect of gender – females evaluated the vernacular variant more harshly than males. Moreover, subjects from South Carolina were found to be more tolerant of *-in* use than subjects from Philadelphia. In the experiment manipulating speaker dialect, they found no differences in overall evaluation, suggesting that the frequency of the apical variant alone was the determining factor in eliciting these overt attitudes.

Campbell-Kibler (2007, 2009) conducted a single-variable study on (ING) using a novel variation of the MGT. Rather than have speech actors read written scripts, Campbell-Kibler recorded hour-long informal interviews then identified stretches of natural discourse with sufficient (ING) tokens. The speakers were brought back into the recording studio and instructed to produce two versions of the passage (one with apicals and one with velars) while mimicking the intonation and rate of the original recording. In contrast to typical MGT manipulations in which the content of the speech remains the same, the use of spontaneous speech does not allow the content to be controlled. Importantly, however, this technique does allow a unique

perspective on the complex interactions between speech style and message content.

Findings revealed a complex interplay between (ING), social perceptions, and attitude. The apical *–in* variant increased the perceived "Southerness" of most speakers, but it also reduced the perceived "gay" accent of another speaker. The velar *–ing* variant also interacted with the perceived socioeconomic status and education of the speaker; for listeners who believed the speaker to be from a working-class background, *-ing* increased the perceived intelligence/education of the speaker. These findings reconfirm many of the results of traditional production studies of (ING). Importantly, however, they also shed light on the flexible nature of perception by showing how a single variant can realize multiple social meanings.

Although MGT studies are pervasive in the field of sociolinguistics, the validity of the technique remains questioned. Lee (1971) has argued that repeating the message forces subjects to focus on form in a manner that may not be commensurate with more naturalistic processing. Gardner and Lambert (1972) acknowledge that the measured attitudes may not accurately reflect underlying mental states, but rather what the participants *think* they should say. Moreover, it is possible that subjects have underlying attitudes which they might not be consciously aware of, nor have overt access to (Greenwald, et al., 2002). Another limitation of the MGT is that it is an offline measure off attitude. That is to say, there is a temporal lag between the presentation of the linguistic stimulus and the attitudinal behavioral response. As such, the semantic differentials used in MGT studies, tap more into deliberative processes rather than more instantaneous, implicit associations.

Implicit Association Test

One way researchers are attempting to more tightly couple the stimulus and attitudinal response

is by using real-time measures. Labov (2008), for example, has run a version of his newscaster experiments using a computer interface that allows participants to register their moment-by-moment attitudinal response by sliding a knob back and forth between two attitudinal poles. One promising technique that we[2] have used in our laboratory is the Implicit Association Test (IAT) (Greenwald, McGhee, & Schwartz, 1998). Because the IAT uses response latency to measure the implicit strength of association between concepts, it circumvents the limitations of other measures of language attitude which require explicit reporting.

During IAT experiments, subjects press buttons in response to the words they read and hear. The difference in mean response latencies measures the strength of the association and its automatic evaluation (Greenwald, et al., 2002). In our sociolinguistic IAT experiments (Table 7-1), subjects are randomly presented spoken words (that end in *–ing* or *–in*) and visual words (associated with our response categories Intelligence vs. Stupidity). If a strong association exists between velar *–ing* and the concept intelligence, then subjects will respond faster when the sorting categories are ING or Intelligence versus IN' or Stupidity (block 4) than when the sorting categories are IN' or Intelligence versus ING or Stupidity (block 7). The strength of association, the D score, is calculated by taking the mean response latency difference between the congruent and incongruent mappings.

Table 3.1 - Illustration of IAT to measure implicit associations between (ING) and Intelligence

Block	No. of trials	Function	Items assigned to left-key response	Items assigned to right-key response
1	20	Practice	Intelligence	Stupidity

[2] Cognitive Neurolinguistics Laboratory (P.I. David P. Corina), Center for Mind and Brain, University of California, Davis.

2	20	Practice	ING	IN’
3	20	Practice	ING or Intelligence	IN’ or Stupidity
4	20	Test	ING or Intelligence	IN’ or Stupidity
5	20	Practice	Stupidity	Intelligence
6	20	Practice	ING or Stupidity	IN’ or Intelligence
7	20	Test	ING or Stupidity	IN’ or Intelligence

We have run undergraduate participants on several different IAT measures (i.e. Intelligence, Gender, and Socioeconomic Status). Results show weak to strong associations between Intelligence and *–ing*, female and *–ing*, and high SES and *–ing*. These results are largely consistent with previous research on (ING) attitude and production. In terms of intelligence, people who use the apical variant are described as *uneducated* and *ignorant*. In regards to socioeconomic status, apical use elicits descriptions that are socially lower (e.g., *working class*, *blue collar*) (Wald & Shopen, 1985), the velar form is associated with higher status speakers and production studies show that apical use is indeed greater for lower SES speakers (e.g., Labov, 1966). In terms of gender, women have a stated preference for *–ing* (Wald & Shopen, 1985) and production studies show that women do indeed use the velar variant in greater frequency than men. The consistency of these preliminary IAT findings with the established literature shows promise for using the IAT in sociolinguistic inquiry.

Importantly, however, IAT results can differ from explicit measure of attitude. Pantos (2010), for example, investigated attitudes towards US- and foreign-accented speech to determine its impact on speaker credibility in a mock medical malpractice trial. Two different measures of attitude were employed: an explicit semantic differential and an auditory IAT. For

the explicit task, participants listened to recordings of US- and foreign-accented speech in the mock trial context. After each physician's testimony, participants rated the speaker on a number of traits (e.g., believability, expertise, etc.) and rated the "fairness" of the verdict. Results revealed an explicit bias in favor of the foreign-accented speaker.

For the IAT, participants heard short auditory snippets (e.g., "it is my opinion"; "I have frequently encountered") and categorized visually presented words that were Good (e.g., marvelous, superb, etc.) or Bad (tragic, horrible, etc.). These results indicated an implicit bias in favor of the US-accented speaker, showing that during automatic evaluative processing, participants are biased towards US-accented speech. One interpretation of these diverging results suggests that participants may have used strategic responding on the explicit measure in order to create a socially desirable persona (e.g., by asserting they are not biased against foreigners). In contrast, implicit measures such as the IAT, are argued to be relatively immune from strategic responding. The contrasting results of Pantos (2010) highlight the dual nature of language attitude and give credence to using multiple measures in experimental design.

3.5 Newly Emerging Methods

In recent years, a number of studies have adopted techniques from cognitive psychology in order to address topics of sociolinguistic relevance. Like the IAT, these techniques allow a tighter coupling between stimulus and response, providing real-time measures of language processing.

Eye tracking measures

Two studies have successfully used eye tracking measures to investigate perceptual processing of dialect and age-graded phonological variation. Eye tracking, which involves monitoring

participants' eye movements, is one of the most widely-used techniques in the cognitive sciences. In terms of language processing, eye tracking has traditionally been used in reading paradigms, though it can be adapted to spoken language. The assumption linking subjects' eye fixations to *spoken* word recognition is that as a word temporally unfolds "the probability that the listener's attention will shift to a potential referent of a referring expression [e.g., a picture or a written word] increases with the activation... of its lexical representation" (Tanenhaus & Brown-Schmidt, 2008).

Dahan and colleagues (2008) used a dialect manipulation to adjudicate between two competing accounts of talker accommodation during speech processing. In some dialects of English the vowel /æ/ is raised towards /ɛ/ before voiced velars, so that the word *bag* sounds more like *beg*. Because raising is limited to voiced velars, this reduces the phonetic overlap of minimal pairs such as *bag* and *back*. The primary question is whether prior exposure to raised /æ/ in words like *bag* influences processing of unaffected counterparts (e.g., *back*). Such an effect would demonstrate the influence of speaker-specific dialect representations on the evaluation of standard "unaccented" input, providing strong support for the dynamic adjustment of representations based on context.

Listeners were exposed to /g/ final words with either a standard /æ/ or a raised variant. On the monitor, participants were presented with four words (e.g., *bag*, *back*, *wig*, *wick*) and were asked to click on the word they heard. Eye gaze was used as a dependent measure of listeners' interpretation of [bæ…]. That is to say, in a word such as *back*, what is the degree to which listeners entertain the competing alternative *bag*? Error rates on the same standard *back*-like words showed that participants' exposure to standard *bag*-like words made significantly more errors (17.9%) than participants exposed to dialect *bag*-like words (7.3%). Analysis

showed that starting around 600 ms after vowel onset, the standard pronunciation group fixated significantly longer on competitor words than the group exposed to the raised variant. These results provide strong support that listeners evaluate speech representations that are dynamically adjusted to take into account context.

A study by Koops, Gentry, and Pantos (2008) used eye tracking to investigate whether listeners have tacit knowledge of age-graded variation in the PIN~PEN vowel merger –front high and mid lax vowels /I/ and /ɛ/ are merged before nasals. Although the PIN~PEN merger is a long recognized characteristic of Southern American English, recent evidence suggests that an "unmerger" is occurring in large metropolitan urban areas of the southern United States.

To test whether Houstonians have knowledge of this age distribution (i.e. that older speakers are merged and younger speakers are unmerged), participants were given a forced-choice word identification task. On each trial, a photo of a young, middle-aged, or older female was presented in the center of the screen, surrounded by four word choices including the target (e.g., *rinse*), a PIN or PEN competitor (e.g., *rent*), and two distracters (e.g., *rack* and *rough*). Importantly, the target and the competitor are temporarily ambiguous (up to and including the nasal) when spoken by somebody with a merged system (i.e. an older speaker). In contrast, the target and competitor are disambiguated at the vowel for un-merged speakers (i.e. younger speakers).

In the experiment, participants listened to words (e.g., "rinse") and clicked on their response as quickly as possible. Of interest was the length of time subjects fixated on the PIN or PEN competitor. If subjects have tacit knowledge of the social distribution of the merger, then PIN or PEN competitors should get more fixations for older speakers compared to younger speakers. Indeed, Koops et al. (2008) found that listeners assume a merged system for "old" compared to

"middle-aged" speakers. However, no significant differences were found between "middle-aged" and "young" speakers, a finding at odds with the production literature that shows that middle-aged speakers align with old speakers.

Event-related Potentials

For the past 30 years, psycholinguists have successfully used electrophysiological measures of brain activity to investigate various aspects of language processing. Event-related potentials (ERPs) are averaged, scalp-recorded measurements of brain activity that are time-locked to the onset of a critical stimulus, such as the last word of a sentence. Because of their excellent temporal resolution (i.e. millisecond processing) and the fact that ERPs can be recorded in the absence of an overt behavioral response, the technique has been widely used to characterize the cognitive processes in language perception and comprehension (Luck, 2005).

A study by Conrey, Potts, and Niedzielski (2005) used the ERP technique to investigate cognitive processing of the PIN~PEN vowel merger in two different populations of speakers: those with merged vowels and those with unmerged vowels. Prior behavioral studies have reported that speakers with merged vowels in production are less able to perceptually discriminate these vowels during perception (Labov, Karan, & Miller, 1991). In the PIN~PEN ERP study, subjects read visually presented sentences that concluded with an expected terminal word, for example, "Sign the check with a *pen*." Following the presentation of the terminal word, a synthesized voice spoke a critical word that was either congruent (/pɛn/) or incongruent (/pIn/) with the visually presented word. Subjects then responded via button press to indicate whether the word they heard matched the word they read.

As expected, behavioral data revealed that the merged dialect group made significantly

more errors identifying incongruent merger stimuli compared to the unmerged group. The electrophysiological data revealed that the unmerged group showed increased waveform amplitude for incongruent, compared to congruent critical words. In contrast, the merged dialect speakers showed no differences in amplitude for congruent or incongruent words. This pattern of results suggest that the two dialect groups differentially process stimuli at a conscious, decisional stage of processing that requires explicit memory of previously encountered tokens.

3.6 Summary

Less than a decade ago, Erik Thomas (2002) lamented that perception was the "neglected stepsister of production in sociolinguistics". In the following years this statement, research investigating the perception of sociolinguistic variation has slowly but assuredly emerged. In this chapter, I have reviewed a number of studies in sociolinguistic cognition which are beginning to cast light on issues that are fundamental to sociolinguistics, including: the relationship between perception and production, how dialect exposure influences processing, how dialects are learned, the nature of sociolinguistic stereotypes, and how social stereotypes influence perceptual processes.

4. Models of Language Processing: Evidence from the N400

In psycholinguistics, an important area of debate concerns *when* and *how* information conveyed through lexical-semantics and larger contexts is integrated during on-line language processing. Two large classes of models have emerged from this debate. One of these classes, two-step models of language interpretation, honors the distinction between semantics and pragmatics and proposes an initial process of context-free sentence-level interpretation (e.g., Lattner & Friederici, 2003). During this stage, the meanings of individual lexical items are integrated with meaning conveyed through morphosyntax (e.g. agent/patient information conveyed by word order or case marking). After the initial sentence-level interpretation has been computed, an additional step has been posited in which sentence meaning is integrated with the wider context. During this step, contextual global factors, such as discourse, pragmatics, world knowledge, speaker identity, and extralinguistic information, play a more prominent role in constraining the sentence-level interpretation. Lattner and Friederici (2003), for example, have argued that mismatches of speaker and message (e.g., a male saying "I like to wear lipstick") are detected after sentence level integration processes. As illustrated in Fig. 4.1, this box plot diagram shows that utterance interpretation (syntactic analysis and thematic processing) is followed by a processing step that integrates sentence-meaning with the discourse model. In this final step, "beliefs about the speaker, and about the speaker's knowledge… as well as a range of

sociolinguistic factors" (p.125), constrain utterance interpretation. In distinction, single-step models of language interpretation propose that a single, unitary process simultaneously integrates both global and local factors during language processing.

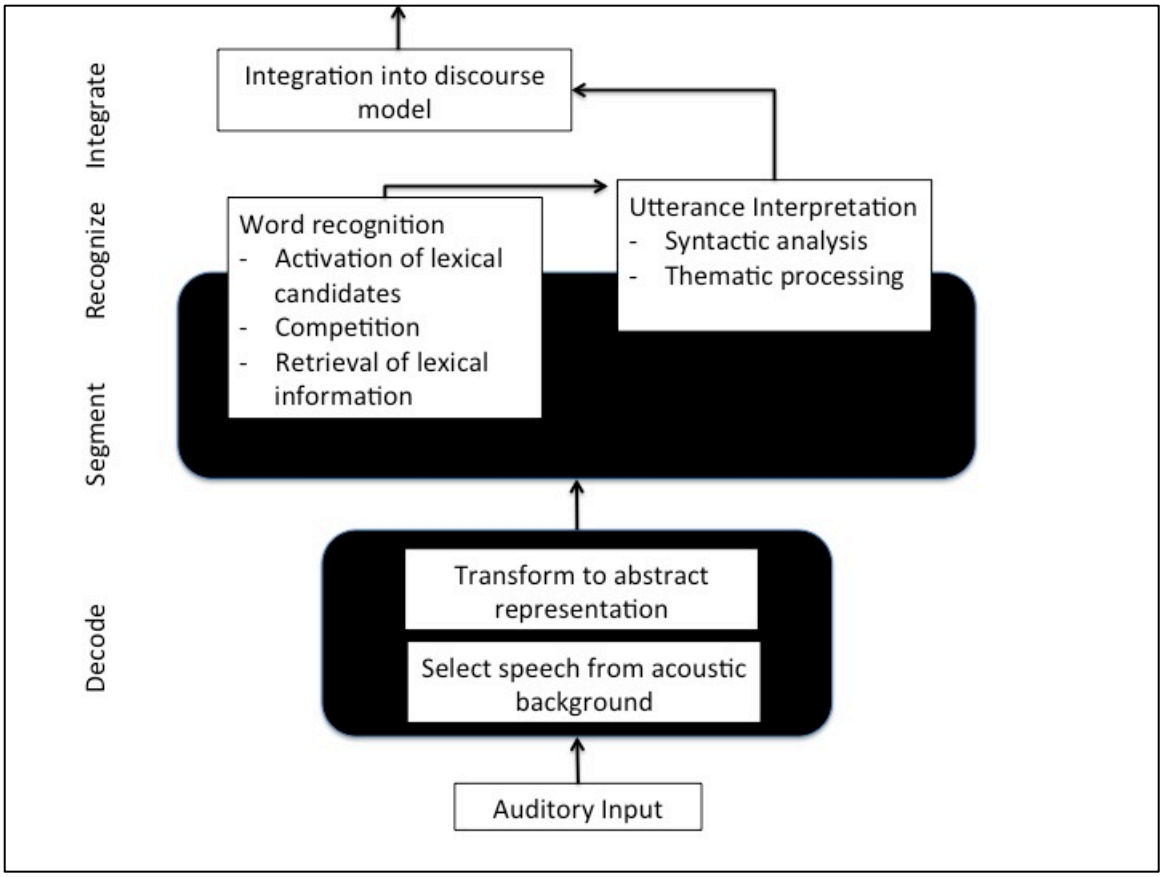

Figure 4.1 - Serial, bottom-up, dual step model of language processing. Adapted from Cutler and Clifton (1999).

4.1 Event-Related Brain Potentials

Because of the rapid nature of language interpretation, experiments that rely on event-related potentials provide an excellent testing ground by which to assess these competing hypotheses. Event-related potentials (ERPs) are measured through electroencephalography (EEG) – scalp-recorded brain activity tracked over time – a technique that provides millisecond resolution of brain processes. Because electroencephalography records the summed activity of cognitive processes (as well as muscular artifacts and line noise), ERP experiments require sufficient numbers of trials per condition in order to increase the signal-to-noise ratio of the raw EEG. The end-result of EEG processing is an array of data reflecting voltage changes across time as recorded at each electrode. ERPs are typically displayed as waveforms consisting of a series of positive and negative voltage deflections, called components. Luck (2005) defines an ERP component as "scalp-recorded neural activity that is generated in a given neuroanatomical module when a specific computational operation is performed" (p.59). Over the past decades, researchers have identified a number of distinct ERP components that are believed to reflect different underlying cognitive processes.

Figure 4.2, reproduced from Luck (2005), shows a typical ERP setup for an "oddball" experiment in which subjects view a series of frequent Xs and infrequent O visual targets and press a button in response to the stimulus. During the experiment, the subject's raw EEG is recorded via an electrode cap that is then amplified, digitized, and saved to a hard drive. The presentation computer the subject is interacting with sends in real-time digital time-lock cues at stimulus onset (i.e. presentation of an X or O). Figure 4.2b illustrates hypothetical data from six trials collected at a single electrode site (Pz). During the averaging process (Fig. 4.2c), an epoch time window is determined (in this case 800 ms) and all the EEG recorded during trials for a

single condition are averaged together, resulting in aggregate waveforms. Figure 4.2c (right) shows the aggregate waveforms for the X and O visual stimuli.

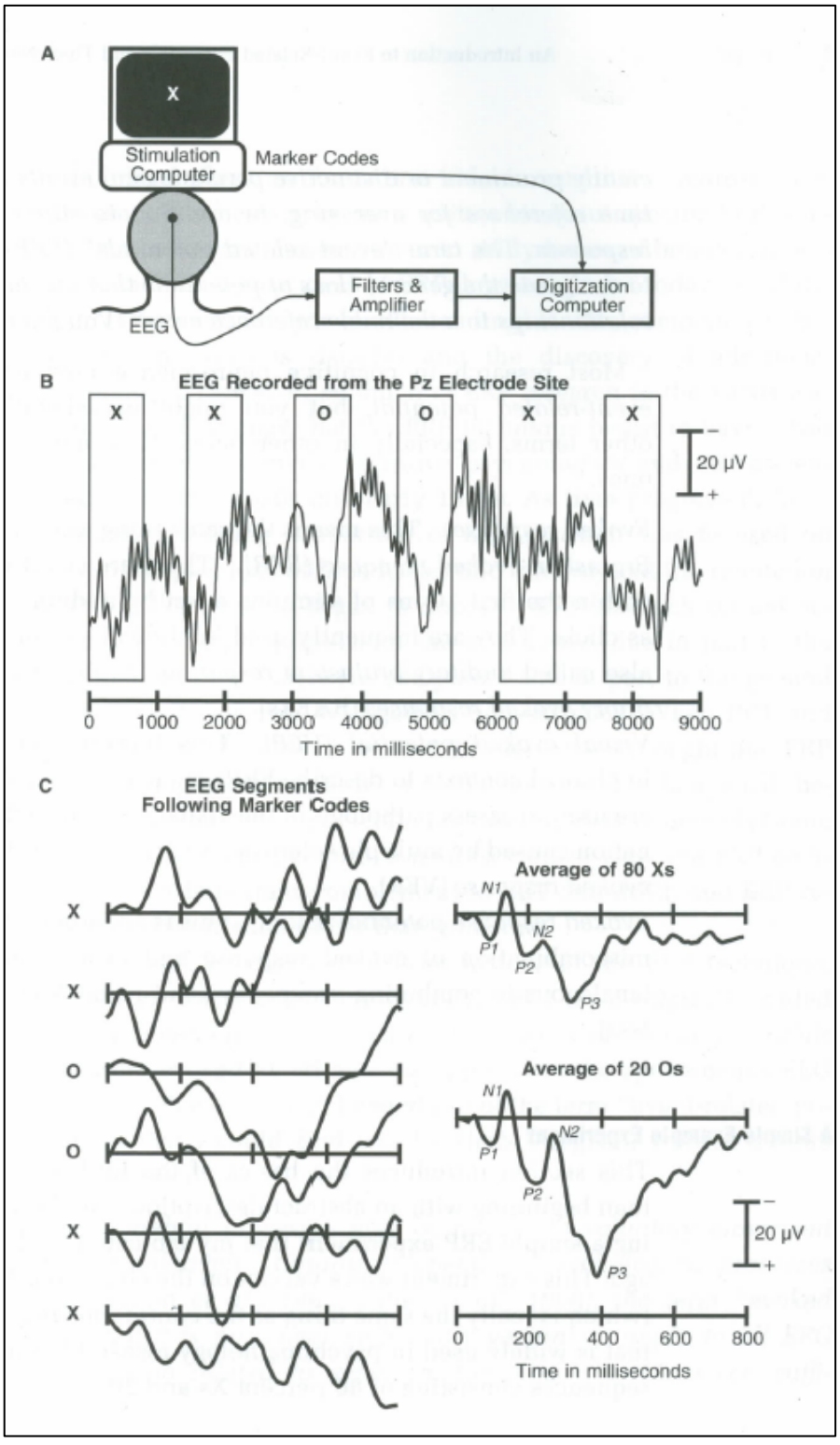

Figure 4.2 - ERP recording and averaging. Negative plotted up. Reproduced from Luck (2005).

4.2 The N400 ERP Component

Because ERP measures have exquisite temporal resolution (Luck, 2005) and do not require an overt behavioral response (Kutas & Delong, 2008), they provide an excellent technique for characterizing the cognitive processes involved in language comprehension. Experiments over the past decades have identified different ERP components which have been interpreted as reflecting different aspects of language processing, such as semantic integration and syntactic parsing (Martin-Loeches, et al., 2009). One well-studied component is the N400, an increased negativity in the waveform that reaches maximum amplitude around 400 ms after stimulus onset and which is typically maximal over centro-parietal electrode sites (Kutas & Hillyard, 1984). In language processing, the N400 is elicited for both visually (Kutas & Hillyard, 1980) and auditorily presented words (Holcomb & Neville, 1991) in sentential contexts. Although all content words elicit an N400 response, component amplitude is modulated by a number of factors including off-line word predictability (i.e. cloze probability), word frequency, and context congruency. Consider, for example, the following sentences (Table 4.1).

Table 4.1- Examples of stimuli used in Kutas and Hillyard (1980).

Condition	**Sentence**	**Relative N400 Amplitude**
Congruent	I shaved off my mustache and *beard*	Small
Improbable	He planted string beans in his *car*	Medium
Anomalous	I take coffee with cream and *dog*	Large

Although each of the sentence final words will elicit an N400 response, component amplitude is function of how well the word "fits" with the sentential context. The congruent condition will elicit the smallest N400 effect; the improbable word will elicit a larger negativity; and the anomalous word will elicit the largest N400 effect. In addition to semantic anomaly, larger N400 responses are also observed for words that are less expected or predicted from the preceding context – "I like my coffee with cream and *sweetener*" (Hagoort & Brown, 1994; Kutas & Hillyard, 1984). Kutas and Federmeier (2011) further note that N400 amplitude is highly correlated with cloze probability – the percentage of subjects who continue a sentence fragment with a specific word.

Recently, a number of studies have demonstrated that the N400 response can be modulated by global contextual factors such as discourse, pragmatics, world knowledge, speaker identity, and extra-linguistic information. Hagoort et al. (2004) reported an N400-like response for world knowledge violations. Ozyurek et al. (2007) found an N400 modulation for violations of co-speech gestures, while a study by Willems et al. (2008) reported similar findings for sentences accompanied with a mismatched picture. Additional studies have reported N400 effects for discourse-level manipulations (Camblin, Gordon, & Swaab, 2007; Nieuwland & Van Berkum, 2006; St. George, Mannes, & Hoffman, 1994; Van Berkum, Hagoort, & Brown, 1999; Van Berkum, Zwitserlook, Hagoort, & Brown, 2003). Importantly, a recent study by Van

Berkum et al. (2008), manipulated voice-inferred stereotypical information (e.g., age, gender, accent) to produce sentences such as "I recently had a check-up at the *gynecologist* in the hospital" spoken by a male voice. The authors report that these violations produced an ERP response that shared the same topography and latency (i.e. temporal delay for stimulus onset), but was smaller in amplitude than the classic N400 elicited by semantic anomaly. Taken together with studies in non-linguistic domains, these ERP studies are beginning to change the "view of the N400 as a marker of language processing" to a view of the N400 "as reflecting meaning processing more broadly" (Kutas & Federmeier, 2011, p. 636).

4.3 N400 Studies informing the Single vs. Dual Step debate

In addressing the issue of single versus dual step language processing, N400 studies have informed this debate. In general, two-step models predict delayed effects for global manipulations such as discourse context, pragmatic, world knowledge, speaker identity, and extra-linguistic information. In terms of N400 modulation, the two-step model predicts: (a) that two distinct ERP components would index the different, underlying cognitive processes of each step in language interpretation and/or (b) a latency modulation of a single component (Van Berkum, et al., 2008). Critically, however, if an effect for global factors is found, it must occur *after* an effect for local factors. That is to say the ERP component or modulation that reflects global factors must have a longer latency than the component that reflects sentence-level

interpretation. In distinction, the single-step model predicts that both local and global factors should show their effects on a single ERP component. In other words, if the latency and distribution of the ERP components are the same for both local and global information, this would provide evidence that a single cognitive process subserves the integration of different levels of meaning. Alternatively, a lack of latency difference between local and global information may indicate that electrophysiological measures are insensitive to multiplexed events. The present paper reviews a number of ERP experiments that manipulate discourse, pragmatics, world knowledge, speaker identity, and extralinguistic information, in order to assess single- and two-step models of language processing.

One question that addresses the concern at hand is when and how *word meaning* and *world knowledge* influence sentence interpretation. Consider the distinctions between the following two sentences: 1.) President Obama is divorced 2.) The house of President Obama is divorced. Sentence #1 has a plausible, coherent interpretation, but the underlying proposition is false in light of our knowledge that president Obama is, in fact, married. Sentence #2, however, has no coherent semantic interpretation because the constraints of the predicate "is divorced" require an animate argument (i.e. houses can't be divorced). Thus sentence #1 is considered a violation of our world knowledge and sentence #2 is considered a violation of some aspect of word meaning. Two-step models of sentence interpretation propose that the interpretation of word meaning and semantics (e.g. animacy constraints) must *precede* the integration of world knowledge or pragmatic information. Single-step models, on the other hand, suggest that both types of information would be processed concurrently.

A study by Hagoort et al. (2004) speaks to this issue. In this study, the researchers used ERPs and fMRI to measure the effects of sentence processing across three different conditions:

pragmatically correct sentences that confirm real world knowledge; sentences that have world knowledge violations, but which are semantically interpretable; and sentences with a semantic violation. Sentences were of the following form: "The Dutch trains are *yellow*/*white*/*sour* and very crowded." Thus, "Dutch trains are yellow" is a correct sentence, because Dutch trains really are yellow. "Dutch trains are white" is a semantically coherent sentence, but one which presents a state of affairs that is counterfactual to the real world knowledge that Dutch trains are, in fact, yellow. "Dutch trains are sour" violates word-level semantics (i.e. only edible things can be described as sour). As illustrated in Fig. 4.3, the researchers observed a classic N400 response for the semantic violation condition compared to the correct sentences. Importantly, they also observed an N400 response to the world knowledge violation which shared a similar distribution and amplitude to the classic N400, and which was identical in onset and peak latency.

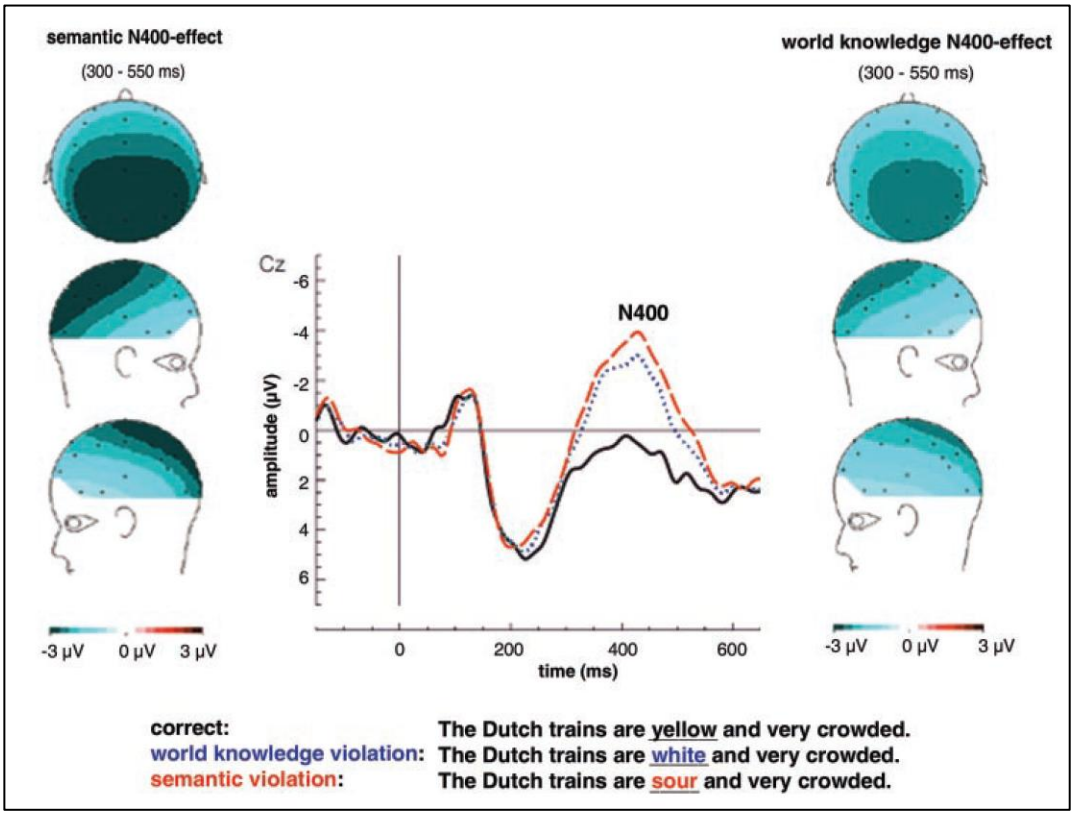

Figure 4.3 - Semantic and world knowledge N400 effect. Negative plotted up. Reproduced from Hagoort et al. (2004).

The same materials used in the ERP experiment were used in an event-related fMRI experiment. Compared to correct sentences, the semantic and world knowledge violation sentences increased activity in left-lateral inferior prefrontal cortex. The results of these two experiments suggest that during sentence comprehension, both word meaning and world knowledge are integrated during the same temporal window and that they recruit the same neuronal networks. These findings are consistent with a single-step model of language interpretation which concurrently processes different levels of information.

Another type of real world knowledge that can be examined is extralinguistic knowledge of speaker identity (e.g., knowledge of the gender, age, and dialect of the speaker. In a recent

study, Van Berkum et al. (2008) examined the effects of speaker identity on language processing. In this study, subjects listened to sentences delivered by different speakers, such that some of the sentences were consistent with the speaker identity ("I recently had a check-up at the *gynecologist* in the hospital" spoken by a female), and some that were inconsistent with the speaker identity ("As we moved house I carried the *washing machine* upstairs" spoken by a female). In addition to gender inconsistencies, subjects also heard sentences that were consistent/inconsistent with regard to age and accent. Since the original discovery of the N400 (Kutas & Hillyard, 1980, 1984), subsequent studies have repeatedly shown that N400 can be systematically modulated by semantic fit of the local context. In order to assess the effects of their manipulation, Van Berkum and colleagues additionally presented subjects with sentences that were anomalous (e.g., "Dutch trains are *sour* and blue") as well as matched control sentences that were semantically correct (e.g., "Dutch trains are *yellow* and blue"). These additional conditions were utilized in order to elicit a classic N400 effect by which they could compare any effects of their speaker consistency manipulation.

As illustrated in Fig. 4.4, the researchers reported an N400 component with greater amplitude for inconsistent versus consistent speaker trials. Though the amplitude of the evoked component was approximately four times smaller than the N400 in the anomalous sentence condition, it shares the same time course (between 200 and 700 ms) and maximum (electrode Pz) as the classic N400 effect. Van Berkum et al., interpret these similarities in time course, polarity, and distribution as evidence that speaker consistency information is processed in the same time frame as local semantic integration processes. In their own words, "voice-inferred information about the speaker is taken into account by the same early language interpretation mechanisms that construct 'sentence-internal' meaning based on just the words" (p. 586).

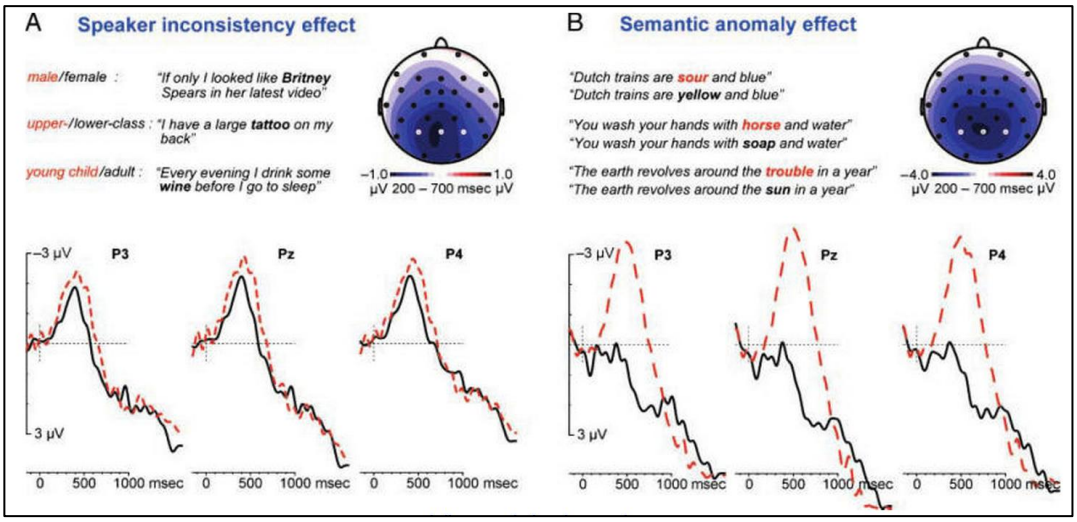

Figure 4.4 - Speaker inconsistency and semantic anomaly effects. Reproduced from van Berkum et al. (2008).

The findings of Lattner and Friederici (2003), are partially at odds with the results reported by Van Berkum and colleagues. Lattner & Friederici, who manipulated speaker gender with stereotypical utterances (e.g., "I like to play soccer"), found no modulation of the N400; rather, they report a late posterior positivity between 600-1000 ms. They take this evidence in support of a two-step model of language interpretation in which properties of the speaker's voice are integrated *after* semantic processing. There are several possible explanations for these divergent findings. Whereas Van Berkum et al., relied on a sentence-medial critical word, Lattner and Friederici relied on a sentence-final word, thus opening up the possibility of sentence wrap up effects. In addition, Lattner and Friederici used short sentences with a single gender manipulation and did not use filler sentences. As Van Berkum and colleagues note, it is possible that subjects discovered the manipulation and attended to the stimuli in a different manner. What is interesting, however, is that Van Berkum et al. also reported a late posterior positivity for their

gender manipulation (but not for their age and accent manipulations). This suggests that additional late processes might be recruited during the processing of gender anomalous discourse.

An additional contextual factor that has been examined in order to assess the validity of various single- and two-step models is co-speech gesture. In natural discourse, speech is often accompanied by extralinguistic information such as gesture. In American English, for example, if somebody is speaking about their crazy friend, the speech utterance "My friend John is a bit crazy" might be accompanied by a "crazy" gesture (i.e. the index finger is moved in a clockwise direction while pointing to the side of the speaker's head). Importantly, however, the co-speech gesture is optional, and the utterance can be interpreted in its absence. A recent study by Ozyurek et al. (2007) used ERPs to investigate how co-speech gestures are processed during sentence interpretation. In their study, subjects listened to speech while watching a video of a person making a co-speech gesture accompanying a critical word. Four different conditions were used: 1.) language gesture match (L+G+) – "He slips on the roof and rolls down" accompanied with a [roll down] gesture; 2.) language mismatch (G+L-) – "He slips on the roof and walks to the other side" accompanied with the [roll down] gesture; 3.) gesture mismatch (G-L+) – "He slips on the roof and rolls down" accompanied with the [walk across] gesture; and 4.) double mismatch (G-L-) – "He slips on the roof and walks to the other side" accompanied with the [walk across] gesture.

The verb in the language mismatch condition, though semantically plausible, is less predictable than the verb in the language match condition. Because semantically less predictable words have been shown increase the N400 amplitude (Hagoort & Brown, 1994; Kutas & Hillyard, 1984), the researchers reasoned this would provide an adequate baseline by which to

compare the latency, amplitude, and distribution of any effects observed for the accompanying co-gesture. In the standard N400 time-window, the researchers found negative deflections for all three mismatch conditions compared to the language gesture match condition. In the mismatched conditions, the components peaked at the same time and shared maximal effects at anterior sites. Further comparisons revealed no significant differences in effect between the three mismatched conditions. The authors speculate that the more anterior distribution (compared to the classic N400) of their mismatched conditions may be attributed to visual processing. The N300 component which is elicited for unrelated pictures compared to related pictures, for example, has this biphasic morphology with a more anterior distribution. . Regardless of this precise interpretation, however, the overall similarity in latency, duration, and relative amplitude – primary indicators of the N400 are consistent with the proposal that the brain integrates both lexical and gestural information simultaneously.

A similar study (Willems, et al., 2008) used ERPs and fMRI to investigate the integration of meaning conveyed through words and pictures. In this study, subjects listened to sentences in which the critical word fit or did not fit the sentential context (e.g., "the man gave his wife a nice *flower*/*cherry* that evening"). Critical words were accompanied by a visually presented line drawing that either matched or mismatched the sentence (e.g. picture of a *flower*/*cherry*). In the ERP study, the double mismatch condition, compared to the double match condition, elicited an N400 with a large negative deflection. Importantly, however, there were no significant differences among the three mismatch conditions, suggesting a common mechanism of semantic integration. In their fMRI experiment, the researchers report increased left inferior frontal gyrus activation for all three mismatch conditions compared to the double match condition. The authors take these as evidence to suggest that despite modality differences, information conveyed

through pictures and words is integrated into sentence context via similar mechanisms, which recruit left inferior frontal cortex during a process of semantic unification. These findings are largely consistent with the previously discussed findings of processing co-speech gestures (Ozyurek, et al., 2007) , though caution is warranted given the modality differences as well differences in distribution and latency which may indicate that partially distinct processes are involved (Kutas & Federmeier, 2011).

The effects of global and local factors on sentence processing can also be examined from the perspective of discourse. One pertinent question for the matter at hand is when discourse-level processing occurs relative to sentence-level interpretation. A study by Van Berkum et al., (2003) used ERPs to investigate these two levels of information in language processing. This study replicated an earlier study (Van Berkum, et al., 1999) but instead of using serial visual presentation, the researchers had subjects listen to the spoken version of the stimuli. In their first experiment, subjects listened to sentences like "Jane told her brother that he was exceptionally *quick*/*slow*." The critical word, which is always acceptable in the local sentential context, was manipulated by the preceding context so that it was either coherent or anomalous with the preceding context (e.g. "her brother was exceptionally *slow*" when in fact, the preceding discourse suggested that he was actually quick). In comparison to discourse coherent trials, discourse anomalous trials elicited a broad negative deflection that peaked around 400 ms and was maximal over centro-parietal sites. Moreover, this classic N400 effect was observed regardless if the critical word was sentence final or sentence medial, and thus cannot be attributed to sentence wrap-up effects. In the second experiment, new subjects were auditorially presented the same stimuli as before, except without the preceding context. In addition, subjects listened to semantically coherent and anomalous sentences (which elicited the classic N400).

Stripped of their biasing contexts, critical words that were previously discourse coherent or discourse anomalous in the prior experiment elicited no significant differences by main effect or interaction. Taken together, the results from these experiments suggest that discourse-level processing is subserved by the same cognitive processes as local, sentence-level processing, and that the interpretative contexts for sentences and discourse are functionally identical.

Although the previous study (Van Berkum, et al., 2003) shows that discourse and sentence level processing occur concurrently, one question that remains unaddressed is whether prior discourse context can override local processing. Modular two-step theories of processing would suggest that local, sentence-level violations should always disrupt the process of sentence interpretation. On the other hand, single-step integrative models of processing might propose that global factors can overrule local interpretation. To this end, Nieuwland and Van Berkum (2006) ran several experiments to investigate whether local animacy violations could be overruled by preceding discourse. Stimuli were humorous short stories composed of six sentences, which were manipulated so that the primary actor was either a man (animate condition) or an inanimate object (inanimate condition). Moreover, sentences contained a transitive verb that formally requires an animate argument (Table 4.2).

Table 4.2 - Example stimulus from experiment #1 (Nieuwland and Van Berkum, 2006).

#	Sentence
1	Once upon a time, a psychotherapist was consulted in her home office by a *yacht/sailor* with emotional problems.
2	The *yacht/sailor* confided in her that everything in life had gone wrong and started crying.
3	The psychotherapist consoled the *yacht/sailor* by stating that everybody experiences these kinds of trouble every now and then.

4	But the *yacht/sailor* doubted whether to continue outlining his problems to her.
5	The psychotherapist advised the *yacht/sailor* to be honest not only with her, but especially with himself.
6	At that moment the *yacht/sailor* cried out that he was absolutely terrified of water

The researchers' primary prediction was that they should observe an N400 response in the inanimate condition, but that the response should attenuate as subsequent sentences are processed. Indeed, for the inanimate condition compared to the animate condition, the researchers observed a significant N400 effect for sentence #1; a trendwise significant diminished N400 in sentence #3; and no significant differences by sentence #5. These findings suggest that prior discourse can neutralize the effects of local animacy violations. In their second experiment, the researchers examined whether the joint effects of real-world plausibility and animacy could be completely overruled by prior discourse. In this experiment, all of the stories (Table 4.3) were about inanimate characters (e.g. a peanut).

Table 4.3 – Example stimulus from experiment #2 (Nieuwland and Van Berkum, 2006).

#	Sentence
1	A woman saw a dancing peanut who had a big smile on his face.
2	The peanut was singing about a girl he had just met.
3	And judging from the song, the peanut was totally crazy about her.
4	The woman thought it was really cute to see the peanut singing and dancing like that.
5	The peanut was *salted/in love*, and by the sound of it, this was definitely mutual.
6	He was seeing a little almond.

The key manipulation occurred in sentence #5, in which the inanimate character was described with a canonically acceptable, but contextually inappropriate property (inanimate condition - the peanut was salted) or the inanimate character was described with a formerly anomalous, but contextually appropriate property (animate condition - the peanut was in love). Grand averaged waveforms revealed a significant N400 response for the inanimate condition

compared to the animate condition, showing that the combined effects of real-world plausibility and animacy can be overruled by contextual plausibility.

The discourse-related findings (Nieuwland & Van Berkum, 2006; Van Berkum, et al., 1999; Van Berkum, et al., 2003) previously discussed are consistent with other studies that have manipulated discourse. A pioneering study by St. George et al. (St. George, et al., 1994), for example, presented subjects with non-coherent passages that could either be preceded or not preceded by a title, the former of which served to make the passage relatively more coherent. Relative to titled passages, the content words of untitled passages elicited a greater-amplitude N400. A study by Camblin et al. (2007) found early (150-250 ms) and persistent (500-900 ms) effects for discourse congruence. Importantly, manipulations of local association did not show effects until the 300-500 ms time window. This is to say, the effects of discourse congruence emerged *before* the effects of association. An earlier study by Van Berkum and colleagues (Van Berkum, et al., 1999), which presented words using serial visual presentation, also reported larger N400 effects for discourse-dependent anomalies. The overall consistency of these findings suggests that discourse representations are integrated concurrently during sentence-level processing. However, questions remain: (1) is discourse-level information always processed during lexical-semantic processing? (2) if not, under what linguistic conditions does discourse processing occur before and/or after semantic processes?

4.4 Summary of the N400 and Language Processing

In summary, this chapter has reviewed a number of ERP studies that manipulated global and local factors to assess two different models of language processing – a two-step model that

proposes local, context-free sentence-level interpretation precedes integration with global factors; and a single-step model which proposes a unitary process that concurrently integrates multiple sources of information during language interpretation. The studies reviewed in this chapter, which examined a wide variety of global factors such as discourse, world knowledge, and extralinguistic information suggest there may be few differences in the way meaning is computed across these varying domains. These studies reported neither a post-N400 component nor a latency-modulation of the N400, which might be taken as evidence of a second step in language interpretation. Rather, these studies suggest that global and local factors under certain conditions can be processed concurrently and that, in some cases, strong contextual effects can even override local processing. Taken together, the most consistent explanation of these findings is a single-step model in which several levels of language representation are simultaneously processed during language interpretation. Importantly, however, the lack of discernible latency differences may alternatively suggest that electrophysiological measures might be insensitive or lack temporal precision in determining the onset of cognitive functions in language processing. Moreover, given experimental differences in modality as well as ERP differences in latency and distribution, suggest that different neural resources may be recruited during language processing (Kutas & Federmeier, 2011).

5. Perception of Sociolinguistic Variation: Evidence from Event-related Potentials

5.1 Introduction

In everyday communication, the words we choose and the way we pronounce them can have important and unintended consequences for both the speaker and the message. Consider vice presidential candidate Sarah Palin's "G-dropping" — "How's that hopey-changey stuff *workin'* out for ya?" On the one hand, this style of speech may have helped endear her with "Joe Six Pack," "Joe the Plumber," and "hockey moms" who have felt alienated from the mainstream political elite. On the other hand, the use of stigmatized, vernacular linguistic forms, clashes with expectations of how vice presidential candidates "should" speak, resulting in the sharp media commentary witnessed during the 2008 U.S. presidential campaign (Purnell, Raimy, & Salmons, 2009). Although sociolinguistic variation is ubiquitous across human languages and can have important social consequences, the underlying cognitive architecture that supports the perception and integration of variable word forms during real-time language processing are poorly understood.

Sociolinguistic variables are "alternative ways of saying the same thing" (Fasold, 1990) which function in part, to characterize differences in register – the varieties of language used in particular situations for particular purposes and social formalities (e.g., a presidential speech vs.

a casual conversation). The linguistic characteristics of different speech registers reflect the situational context surrounding the speech event, including the audience, the location, and the topic (Bell, 1984; Ervin-Tripp, 1973). Informal conversational registers, for example, are characterized by increased frequency of vernacular variants (e.g., Douglas-Cowie, 1978; Russell, 1982) , which in American English include increased contraction (Biber, Johansson, Leech, Conrad, & Finegan., 1999), and consonant cluster reduction (Guy, 1980; Labov, 1972). More formal registers, on the other hand, tend to be characterized by a relative reduction of these vernacular variants in favor of canonical forms.

In the present experiment we used event-related potentials (ERPs) to examine the processing consequences associated with speech that varies along a single sociolinguistic dimension – the realization of the English progressive morpheme (ING) which alternates between two pronunciations (e.g., *working/workin'*). The factors constraining the production of this variable have been well-documented across varieties of English spoken across the globe (Douglas-Cowie, 1978; Hazen, 2008; Kiesling, 1998; Labov, 1966, 2001b; Reid, 1978; Roberts, 1994; Shopen, 1978; Shuy, Wolfram, & Riley, 1967; Trudgill, 1974; Wolfram & Christian, 1976). One well-attested finding is that the formality of the conversational context influences which form will be realized in speech production, with more formal social contexts promoting the ING pronunciation and more casual settings promoting IN' realization (Douglas-Cowie, 1978; Labov, 1966; Trudgill, 1974; Wald & Shopen, 1985).

While the factors governing the realization of sociolinguistic variables and their attendant social attitudes have been well documented, it is only recently that researchers have begun to investigate how variable forms such as (ING) are processed during real-time language comprehension.

Much of the work on how socially variable word forms are processed, has relied on offline measures which focus on the perception and categorization of regional vowel variants (for full reviews see Clopper & Pisoni, 2005 and Chapter 3, this dissertation). For example, Plichta and Preston's (2005) categorization study of monophthongal /ay/, a feature of Southern American English. In this study, seven resynthesized vowels were inserted in a [gVd] frame, to create words ranging from a fully diphthongal 'guide' to fully monophthongal 'god.' Participants listened to these words and clicked a map to indicate one of seven locations on a north-south continuum. Results showed that participants can accurately discriminate and geographically categorize these regional vowel variants. A subsequent study by Plichta and Rakerd (2010) showed that participants with exposure to the Northern Cities Chain Shift (NCCS) – a shift of vowels occurring in the Inland north dialect region of the U.S. – compared to those without exposure, shifted their vowel discrimination thresholds when listening to the speech of an NCCS speaker compared to a non-NCCS speaker. Sumner and Samuel (2009), used prime target word pairs in a lexical decision task across three priming experiments, short term form priming, semantic priming, and a blocked long-term repetition priming. Results showed that New York City participants who produce r-less word variants (e.g., "baker" → bak[ə]) in their speech, are equally primed for both rhotic and non-rhotic variants, while speakers with canonical r-full pronunciations (e.g., bak[ɚ]) are only primed for this specific variant. These studies, suggest that listeners are sensitive to these variants, but the methods relied upon are largely off line measures rather than real-time measures of naturalistic discourse. While these studies highlight the importance of linguistic exposure on language processing, a number of recent studies have further illustrated how processing is modulated by extra-linguistic indexical information including sociolinguistic stereotypes.

Niedzielski (1999), for example, investigated the relationship between speech perception and linguistic stereotypes – specifically Canadian Raising, a phonetic process in which the /aw/ diphthong (as in the word *house*) is raised and fronted compared to canonical /a/. In this study, all sentences were spoken by a single speaker with raised /aw/, but importantly, participants categorized the *same* vowel tokens differently depending on whether they *believed* the speaker was from Detroit or a nearby Canadian city. A study by Hay and colleagues (2006), showed similar vowel stereotype effects by priming participants with the word 'Australia' or 'New Zealand' at the top of the answer sheet. When given the Australia answer sheet, listeners perceived more raised and fronted variants than when given the New Zealand answer form. Together, these studies suggest that perceived speaker origin is sufficient to bias phonological processing during certain categorization tasks.

Additional studies in sociolinguistic cognition have further illustrated how language processing is biased by extra-linguistic information such as the perceived age, race, sex, and occupation of the speaker. Strand and Johnson (1996) showed that participants shift their /s/-/ʃ/ boundary when listening to a female compared to a male speaker, or when shown a female photo versus a male photo. A study by Staum (2008) on –t/d deletion showed that perceived speaker race speeds reaction times for ambiguous items when the items are presented with a race photo that is consistent with the sociolinguistic phenomena. Similarly, Koops, Gentry, and Pantos (2008), using an eye-tracking paradigm to investigate the PIN~PEN vowel merger, showed that the perceived age of a speaker influences gaze fixation to competitor words that are partially ambiguous for a speaker with a merged vowel system.

Finally, Labov et al., (2006) has investigated the role of variant frequency in explicit attitudinal response. Using a matched guise technique, subjects listened to short passages read by

a "newscaster in training" and rated these passages on how "professional" the newscaster sounded. Passages varied in frequency of the English (ING) morpheme, being realized with either its canonical velar pronunciation (*e.g., working*) or in its vernacular apical form (e.g., *workin'*). Results showed a logarithmic relationship between vernacular IN' frequency and negative judgments of professionalism (Fig. 2.5). Although these studies illustrate the importance of extra-linguistic indexical information in language processing and language attitude, these methodologies are limited in that they do not do not permit insight as to how and when sociolinguistic variation is processed in real time in the service of language comprehension.

As discussed in Chapter 4, there is a general consensus among psycholinguists that language comprehension involves at least two fundamental operations: accessing representations stored in memory (lexical access) and integrating these structures into higher order representations (integration) (Jackendoff, 2007; Lau, Phillips, & Poeppel, 2008). Lexical access entails an initial mapping of the incoming acoustic signal to the meanings of words stored in long-term memory. Sentence-level meaning accrues from the integration of these individual word meanings with information conveyed by syntax (e.g., word order) (Pinker & Jackendoff, 2005; Radford, 1997). In addition to the local sentence meaning, pragmatic and contextual factors, such as world-knowledge and speaker-knowledge additionally constrain the ultimate interpretation of an utterance (Hagoort, Hald, Bastiaanesen, & Petersson, 2004; Van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008). There is, however, as previously discussed in Chapter 4, some debate as to when factors, such as socially pragmatic information, including sociolinguistic variation, influence sentence interpretation.

In psycholinguistics, two general models have been proposed. The *two-step model* honors the distinction between semantics and pragmatics, and proposes an initial process of context-free

sentence-level interpretation. After an initial sentence-level interpretation has been computed, an additional step integrates sentence meaning with the broader context. During this time, contextual factors, such as discourse, pragmatics, and extra-linguistic information, play a role in constraining the sentence-level interpretation. Lattner and Friederici (2003), for example, have argued that speaker message mismatches (a male saying "I like to wear lipstick") are detected after semantic processing. Similarly, in Cutler & Clifton (1999), their serial box plot diagram shows that utterance interpretation (syntactic analysis and thematic processing) is followed by a processing step which integrates sentence-meaning with the discourse model. In this final step, "beliefs about the speaker, and about the speaker's knowledge… as well as a range of sociolinguistic factors" (p.125), constrain utterance interpretation, though the authors qualify this by noting the lack of data capable of resolving this issue. In contrast, single-step models of language interpretation propose that a single, unitary process simultaneously integrates both contextual and local factors during language processing (Hagoort et al., 2004; Van Berkum et al., 2008). These models argue that social and pragmatic knowledge is integrated during utterance interpretation by the same early systems that integrate the meanings of individual words into larger conceptual representations. For further discussion of these two different processing models, see Hagoort and Van Berkum (2007) and Van Berkum et al. (2008).

Electrophysiological studies assessing N400 effects during language processing have been used to inform this debate. The N400 ERP component is a broad negative deflection of the waveform that peaks 400 ms after the visual or auditory presentation of a word, and is typically maximal over centro-parietal electrode sites (Holcomb & Neville, 1991; Kutas & Hillyard, 1980). Although all content words elicit an N400 component, the ERP response is larger for words that are semantically anomalous or less expected (Hagoort & Brown, 1994; Kutas &

Hillyard, 1984). For example, listening to the sentence 'I like my coffee with cream and *sugar/sweetener/socks*', the expected word *sugar* elicits the smallest N400, the plausible but less expected *sweetener* elicits a larger waveform, and the semantically anomalous *socks* elicits the largest N400 response. Thus the N400 is often interpreted as an index of ease or difficulty in semantic conceptual integration (Brown & Hagoort, 1993; Hagoort & Van Berkum, 2007).

Consistently, studies over the past three decades have shown that the closer a word's meaning fits with the prior context (broadly construed), the greater the reduction in N400 amplitude. Importantly, however, N400-like effects show differences in amplitude distributions and latency depending on modality (visual/aural), stimulus characteristics, and task (Kutas & Federmeier, 2011). Recently, a number of studies have demonstrated that the N400 response can be modulated by specific contextual factors such as discourse, pragmatics, and world knowledge (Camblin, Gordon, & Swaab, 2007; Hagoort et al., 2004; Nieuwland & Van Berkum, 2006; St. George, Mannes, & Hoffman, 1994; Van Berkum, Hagoort, & Brown, 1999; Van Berkum, Zwitserlook, Hagoort, & Brown, 2003). For example, a recent study by van Berkum and colleagues showed that socially indexical factors conveyed by speaker voice modulates the N400 response. In this study, subjects listened to passages read by different speakers whose voice conveyed socially stereotypical information which was either congruent or incongruent with the message (e.g., "Every evening I drink some *wine* before I go to sleep" spoken by an adult or by a child). Results showed a small, but significant N400 effect for incongruent passages spoken in a voice that violated these stereotypical associations. Taken together, these studies have provided additional electrophysiological support for a single-step model of language processing.

In the present study we investigated processing the sociolinguistic properties of speech, specifically the realization of the English progressive morpheme (ING) in formal and informal

register contexts. From a psycholinguistic perspective, the study of (ING) is of particular interest, in that although variable word forms such as *working/workin'* convey the same denotational linguistic meaning, they may additionally convey extra-linguistic, connotational meaning through frequency-based association with different sociolinguistic factors (e.g., race, gender, region, register, etc.). Here, we address whether sociolinguistic variation engenders processing costs that would modulate the N400 response, thus providing insight into the time course and integration of extra-linguistic, socially-indexical meaning.

In our experiment, participants listened to spoken stories that were designed to create high expectations for specific words (cloze probability) and specific sociolinguistic variants (Fig. 5.1). A word's cloze probability is the percentage of individuals who choose that word to complete a sentence fragment in an offline sentence completion task (Delong, Urbach, & Kutas, 2005). In some of the stories, the expected high cloze word form was encountered (e.g., the baseball game was cancelled because it started *raining*), while in others, less expected low cloze forms were encountered (e.g., the baseball game was cancelled because it started *storming*). Half of the stories concluded with canonical ING variants (*raining/storming*), while the other half concluded with vernacular IN' variants (*rainin'/stormin'*). The sentences leading up to the critical word were designed to test how the preceding register context influences expectations for the critical variant. Specifically, these preceding register contexts used verb forms with either the ING pronunciation (formal passages) or the IN' pronunciation (informal passages).

The single and dual-step models of language processing, make different predictions for these manipulations. If word meaning and sociolinguistic meaning recruit the same early sense-making processes, we expect to observe N400-like responses for both word expectancy and sociolinguistic expectancy. If we observe a brain response to sociolinguistic variation within the

temporal confines of a "classic" N400 window, and if that response shares a similar time-course, morphology, and amplitude to the N400, then we would have good evidence that socially indexical meaning and word meaning are integrated concurrently, consistent with a single step model of language processing. In contrast, the two-step model of language processing predicts that any effect of sociolinguistic variation should occur after lexical-semantic processing. In terms of ERPs, the two-step model predicts delayed, possibly different effects compared to a classic N400.

5.2 Predictions

5.2.1 Cloze Probability and Sociolinguistic Variant Predictions

In this study, we chose a fully balanced design, one in which we presented passages in formal and informal registers which had corresponding critical words realized with the ING or IN' variant. Given the previously discussed ERP and sociolinguistic literature while taking into account the relative infrequency of the vernacular IN' variant in this speech community (see Variant Frequency Measures in Methods below), we predicted: (1) similar timing and morphology of N400-like potentials for cloze probability and variant relative to the acoustic onset of the word and variant respectively; (2) increased negativities for low compared to high cloze words; and (3) increased negativities for vernacular IN' variants compared to canonical ING variants.

5.2.2 Congruency & Variant Interaction Predictions

In addition, we entertained two different predictions of the ordering of effects based on our experimental design: a "congruency" account and a "frequency-based" account of sociolinguistic processing. The congruency account, predicts that the preceding register context will strongly determine listener expectations for similar variants (i.e. Formal→ING & Informal→IN'). That is to say, hearing a series of ING words in a discourse context should prime expectations for ING critical words, while hearing a series of IN' words should engender expectations for vernacular IN' variants. This account predicts that incongruent critical variants that mismatch the preceding register context (i.e. Formal→IN' & Informal→ING), should elicit a processing cost reflected by heightened N400 negativities compared to congruent variants. Taking into consideration the relative infrequency of the vernacular IN' variant in this speech community, this account further predicts that IN' critical words will have greater negativities than canonical ING words independent of the word's congruency with the preceding register context. The predicted ordering of effects is displayed in Table 1 (left).

If local register contexts are insufficient to engender sociolinguistic expectations, however, then one might expect the ordering of conditions to reflect the sociolinguistic behavior of the speech community under study. In other words, the less preferred or produced variant (IN') should elicit heightened N400-like negativities compared to the preferred variant (ING), and the ERP response should be a function of *overall* frequency of the less preferred variant in each of the passages (c.f., Labov et al., 2006). This frequency-based account predicts that critical words preceded by an informal register context should in general be more negative than those preceded by a formal context (i.e. Informal < Formal), and that within each of these levels, IN' words should be more negative than ING words. The proposed hierarchy of effects for the frequency-based account is illustrated in Table 5.1 (right).

Table 5.1 - Predicted ordering of effects. In order to facilitate discussion and presentation of results, we rely on the following notation: Register→Variant, where Register = Formal or Informal, and the critical Variant = ING or IN'.

μV	Congruency Account	Frequency Account
More +	Formal→ING	Formal→ING
	Informal→IN'	Formal→IN'
	Informal→ING	Informal→ING
More -	Formal→IN'	Informal→IN'

5.3 Methods

5.3.1 Participants

We recruited 45 undergraduates enrolled in an introductory psychology class at the University of California, Davis. All participants gave informed consent and were given course credit for participating in the experiment. Due to technical issues, four subjects were excluded from analysis; the data from 12 subjects with excessive alpha activity and/or skin potentials, which resulted in fewer than 20 artifact-free trials per condition, were also not included. Data analysis is based on the 28 remaining subjects. These subjects, 14 men (18-23 years, mean age = 20.4 years) and 14 women (18-23 years, mean age 19.3 years) were neurologically unimpaired native speakers of English with normal hearing. 25 participants were right handed and three were left handed. All participants were raised in the western United States (26 California, 1 Nevada, 1 Arizona).

5.3.2 Materials

In this study we crossed target word and variant expectancy with the sociolinguistic register of a preceding passage (Fig. 5.1). Time-locked critical words were sentence-final verbal forms inflected for progressive morphology (using either the canonical velar ING or vernacular apical IN' pronunciation, see below). We manipulated semantic expectancy by preceding the critical words with a short, semantically rich passage. These stimuli were developed so that subjects would have high expectations for the passages to conclude in a specific word (e.g., "…the game was cancelled because it started *raining*"). We established cloze probability using an offline task in which 30 subjects took a pen-and-paper test containing written versions of the stimuli with the final word missing (e.g., "… the game was cancelled because it started _____"). Subjects were asked to fill in the blank with the *-ing* word that best completes the passage. Subjects that participated in the cloze probability task did not participate in the EEG study. Critical words in the high cloze condition had a mean cloze probability of 74.9% (range 20.7% - 100%) whereas low cloze alternatives had a 1.2% mean cloze probability (range 0% - 13.8%).

The passage preceding the critical word was manipulated for sociolinguistic register. Specifically, each passage contained four embedded words inflected with progressive morphology. In formal passages, the canonical ING variant was exclusively used, while in informal passages the vernacular IN' variant was used. In addition to cloze probability, the sentence-final critical words were also manipulated for sociolinguistic variant (taking either the ING or the IN' pronunciation) and register congruency (congruent or incongruent with the preceding passage). In summary, we created a 2x2x2 factorial design using the factors cloze probability (Cloze – high/low), sociolinguistic variant (Variant – ING/IN'), and register

congruency (Congruency – congruent/incongruent). All critical words were unique and never appeared in the fillers or in the passages of other stimuli.

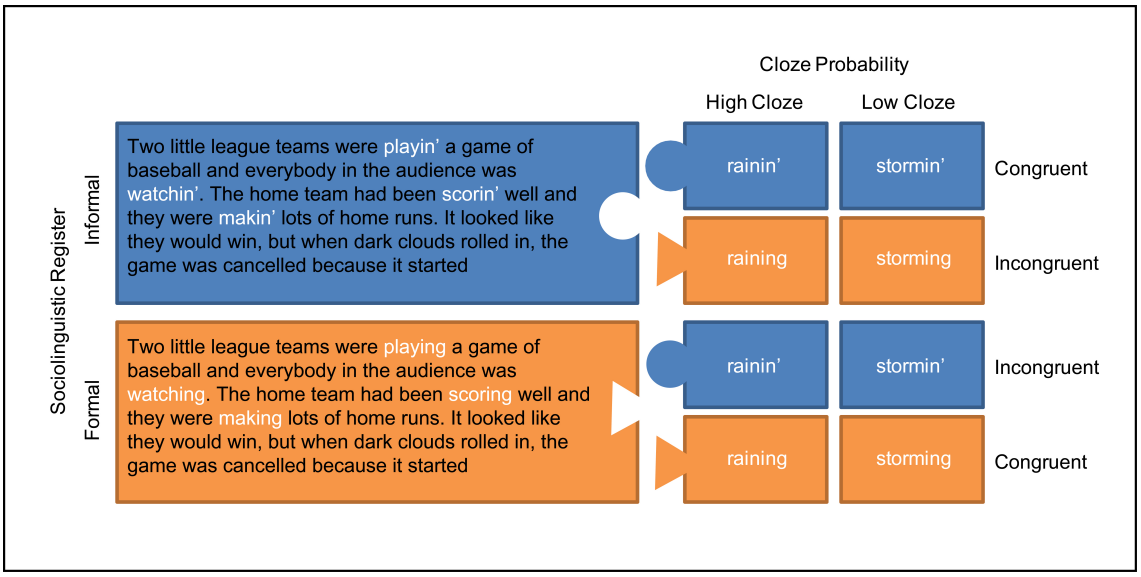

Figure 5.1 - Design of experimental stimuli. Each stimulus consisted of a short passage read by a speaker in either a formal or informal guise. Informal passages (blue) used IN' variants, while formal passages used ING variants (orange). Sentence-final critical words varied in offline predictability (high cloze or low cloze), were spoken using either the IN' (blue) or ING pronunciation (orange), and were either congruent or incongruent with the formality of the preceding passage.

In all, we developed 280 critical passages and randomly assigned them to six native speakers of English (three female). Speakers were all graduate students of linguistics at the University of California at Davis and ranged in age from 24-40 years (mean 31.6 years). Speakers were originally from or had spent substantial time in California, and all spoke a standard variety of Western American English as impressionistically assessed by the first author. The salient phonetic features of these speakers included the fronting of high back vowels (/uw/ 'boot', /ow/ 'boat', and /U/ 'book'), the merging of low back vowels (/ɑ/ 'pot' and /ɔ/ 'bought'), and other features characteristic of the speech of this region (Eckert, 2008).

Passages were read from scripts in two different guises – once with all velar ING pronunciations and once with all apical IN' pronunciations. Speakers were instructed to maintain the same speech rate, pitch, and prosody between guises. All critical words were read separately in a phrasal context in order to maintain natural prosody and coarticulation. The phrasal frame (e.g., the game was cancelled because it started *rainin'*) was then digitally spliced in its entirety into the preceding context (van Alphen & van Berkum, 2009) to produce eight versions of each passage corresponding to the eight conditions. Phrasal frames consisted of the critical ING/IN' word plus 1-15 words immediately preceding it (See appendix A for example stimuli and their phrasal frames). Each stimulus consisted of a single .wav file with one embedded trigger cue at the onset of the sentence-final critical word and one cue at the onset of the ING/IN' allomorph. Eight lists were created such that each list contained 35 stimuli from each of the eight conditions balanced across speakers. Lists were interspersed with 120 filler passages (without any ING or IN' words) read by the same six speakers.

One concern with our design is whether any observed effects reflect sensitivity to factors other than sociolinguistic variability. To address this concern, we examined sentence plausibility and lexical factors, including word identification times for our critical word targets, prosodic variability in these tokens, and the frequency of variant forms in this population.

5.3.4 Sentence Plausibility Task

Following Van Berkum et al. (2008), we recruited 29 additional subjects to participate in a plausibility test, in order to determine whether subjects find IN' sentences less plausible than ING sentences. In this task, subjects listened to spoken sentences and rated them on a five point scale of plausibility according to the instructions: "how normal or strange you think it is to have

the speaker say this particular thing" (1 = completely normal, 5 = very strange). Stimuli consisted of the final sentence from our experimental stimuli (e.g., It looked like they would win, but when dark clouds rolled in, the game was cancelled because it started raining). We included 5 different sentence types: ING – 50 (… *raining/storming*), IN' – 50 (… *rainin'/storimin'*), Semantic Anomaly – 25 (e.g., "I ran out of cat food last week so I fed my cat an *elephant*."), Speaker Violation – 25 (e.g., "When we went to dinner last week, Tom spilled coffee all over my *dress*.", spoken by a male), and Control Sentences – 50 (e.g., "Last weekend, Tom and Cindy went on a date"). Stimuli were counterbalanced across lists and randomly presented. As expected, semantic anomaly sentences were rated less plausible (mean = 4.6, SD = 0.3, range = 3.9-5.0) than control sentences (mean = 1.4, SD = 0.4, range = 1.0 – 2.5). Sentences with speaker anomalies were also rated as less plausible (mean = 3.3, SD = 0.8, range = 1.3 – 4.8). Both ING sentences (mean 1.8, SD = 0.5, range = 1.2 – 2.9) and IN' sentences (mean 2.1, SD = 0.8, range = 1.1 – 3.9) were rated slightly less plausible than control sentences. Importantly, no difference in ING and IN' sentence plausibility was observed [t(56) = -1.71, *n.s.*]. In terms of plausibility, the following ranking was observed: control > (ING and IN') > speaker violation > semantic anomaly.

5.3.5 Gating Task

In order to identify the point at which individuals could reliably identify the critical allomorph, a separate gating study was conducted on a separate pool of subjects (n=45) using 100 ING/IN' pairs (e.g., *raining/rainin'*) randomly selected from the target words used in the ERP study. These words were split into smaller wave files which increased in duration by 50 ms increments (e.g., 0-50 ms, 0-100 ms, 0-150 ms, and so on), until the entire length of the word was captured. This resulted in each word being split into 9-16 smaller sound files. Two lists were

created: List 1 consisted of 50 ING words and 50 different IN' words, while their ING/IN' counterparts were assigned to list 2. Stimuli were randomly presented through a loud speaker, and subjects, who were counterbalanced across left and right key mappings, indicated via button response whether they thought the sound snippet was an ING word or an IN' word. The identification point was determined by calculating the mean gate at which 80% of participants could correctly identify the stimulus, with subsequent larger gates maintaining an 80% or larger correct identification rate through the largest gate. The mean isolation point for ING words was at gate 7.19 (SD 1.77) and that for IN' words was at gate 7.07 (SD 1.66) indicating no systematic differences in ING/IN' identification gate [$t(99) = .50$, $p = .62$]. We used these data to establish the average ING/IN' identification point at approximately 360 ms after word onset.

5.3.6 Acoustic Measures

In order to document the acoustic properties of our stimuli we conducted two additional analyses of prosody and intensity in Praat (Boersma, 2001) on the same 100 critical word pairs used in the gating study. A paired-samples t-test was conducted to compare mean F0 in the two groups of ING and IN' words. It is important to bear in mind that our intent of this study is not to determine the minimal acoustic difference(s) which may signal the presence of, or absence of an ING versus an IN' variant. Rather, we are interested in assessing the consequences of encountering the naturally produced variant forms within and across register contexts. As such we presented the critical targets, which have inherent acoustic differences, as they appear in situ. Our analysis revealed non-significant trends in the mean F0 scores for ING (M=160.32 Hz, SD=54.3) and IN' (M=154.30 Hz, SD=48.91) conditions [$t(99)=1.78$, $p = 0.08$] and intensity

between ING (M=66.11 dB, SD=3.89) and IN' (M=66.46 dB, SD=3.60) conditions [t(99)=-1.91, p = 0.06].

5.3.7 Variant Frequency Measures

In order to assess the prevalence of vernacular word forms in this speech community, we conducted two additional production studies with undergraduate subjects who did not participate in the main EEG experiment. In the first study, sociolinguistic interviews were conducted with 7 native Californian subjects (female=3). Using production data as a proxy for underlying strength of the cognitive representation, we assumed that conversations among friends would provide the most accurate measure of variant frequency as it is encountered by this undergraduate population. Two age-matched female undergraduate research assistants conducted hour-long interviews with their friends on topics of interest to this population (e.g., school, dating, parties, etc.). Three undergraduate students of linguistics transcribed these data and impressionistically coded the ING/IN' variant. These interview data revealed a mean vernacular IN' occurrence of 9.5% (ING – 419, IN' - 44).

In a second production study, we recruited 12 participants (female=6) who were given a video description task. These participants watched two back-to-back 5 minute episodes of an animated cartoon video about a penguin family. While watching the videos, participants were asked to "describe the actions taking place in such detail that somebody listening only to your voice can adequately visualize what is happening in the video." On average, each participant produced 72 verbal forms with progressive inflection (min=33, max=122). The majority of these tokens were realized with the ING pronunciation (ING=796, IN'=58), with the apical IN' pronunciation only accounting for 6.1% of the total. Importantly, however, we did note

differences in vernacular variant production by gender in this limited sample, with males showing greater vernacular usage than females (males: 10.2%, females: 3.3%). Taken together, these production results suggest that vernacular IN' variants are relatively infrequent in this community, with about 90% of progressive morphemes realized with canonical ING pronunciations.

5.3.8 Procedure

After the electrodes had been applied, participants sat in a sound-attenuated chamber and listened to 400 short passages spoken by the six different speakers. Stimuli were presented using Presentation (version 14.4) and were played from a loud speaker positioned ~1.5 meters in front of participants. Participants were asked to listen to passages for comprehension and they were instructed that they would be presented true/false statements about the passages they just heard which would require a button press to continue. A short practice session ensured that participants fully understood the instructions and task.

When each trial began, a black fixation-cross appeared in the middle of the computer screen. Five seconds before the onset of the sentence-final critical word, the fixation cross turned red and remained on the screen until 1500 ms after the sound file finished playing. While the cross was red, participants were instructed that they should refrain from blinking, but they were otherwise encouraged to deliberately blink at all other times. After each trial, there was a 33% chance that a true/false question relating to the passage would be visually presented, requiring a button response. Stimuli were randomly presented across nine blocks of approximately 10-15 minutes each and subjects were allowed a short break between blocks. The total length of the experiment, including preparation and subject debriefing was approximately 2.5 to 3.0 hours.

5.3.9 EEG Recording and Analysis

Participants' electroencephalogram (EEG) was recorded from a 32 channel electrode cap using the BioSemi ActiveTwo system. Six additional external electrodes were used for off-line signal processing: one electrode each at the left and right mastoid, one electrode each at the left and right outer canthi, and one electrode above and one below the left eye. After recording, the EEG was re-referenced offline to the averaged left and right mastoids. The electro-oculogram (EOG) was calculated by the HEOG and VEOG electrodes using an offline conversion to horizontal and vertical bipolar signals. During recording impedances were maintained below 20 kΩ and data was sampled at 256 Hz.

Prior to artifact detection, the EEG signal was filtered off-line with a band pass filter set to .01 and 30 Hz (Luck, 2005). Artifacts were identified using the algorithms provided by the ERPLab plugin to EEGLab. Across the entire epoch for each of the 38 channels, a moving window peak-to-peak (voltage) threshold was set (100 μV) and calculated using a 100 ms wide window taking 50 ms steps (erp_artmwppth). Blink detection was calculated on the VEOG channel, with the blink width set to 350 ms at a normalized cross-covariance of 0.65 (erp_artblink). Saccades were detected on the HEOG channel using a moving window of 200 ms taking 50 ms steps with a threshold of 30 μV (erp_artstep). After automatic artifact detection, trials were double checked by hand to ensure correct identification of artifacts. Segments with artifacts due to eye blink, saccade, or movement (13.6% evenly distributed across conditions) were not included in analysis.

For each subject, an average waveform was computed of artifact-free trials for each of the eight experimental conditions across four electrode regions: anterior left (al) - Fp1, Af3, F7,

F3, Fc1, Fc5; anterior right (ar) - Fc6, Fc2,F4,F8, Af4, Fp2; posterior left (pl) - Cp1, Cp5, P7, P3, Po3, O1; and posterior right (pr) - O2, Po4, P4, P8, Cp6, Cp2. Each epoch of 1200 ms included a 200 ms pre-stimulus baseline followed by a 1,000 ms post-onset window. Epochs were divided into 20 adjacent windows of 50 ms, and the mean amplitude for each 50 ms window was entered into a repeated measure ANOVA which included the following factors: Cloze (high/low), Register (formal/informal), Variant (ING/IN'), AP (anterior/posterior), and LR (left/right). When two or more adjacent 50 ms windows showed the same significant ($p < 0.05$) effects, these smaller windows were collapsed into a larger window in which we determined the onset and offset of effects by statistical analysis using 10ms windows. We report our findings in terms of ANOVAs conducted on these larger, collapsed windows.

5.4 Results

Cloze Effect 230-510 ms (word onset)

As predicted, we observed a broad negativity peaking around 400 ms with a posterior distribution. The main effect of Cloze became significant at 230 ms and lasted until 510 ms [Cloze: $F(1,27) = 17.06$, MSE 8.37, $p < .001$]. As illustrated in Fig. 5.2, we observed a Cloze x AP interaction, with low cloze words (dashed black) showed significantly greater negativities than high cloze words (solid red) at posterior electrode sites [-1.341 μV difference, $F(1,27) = 24.97$, MSE = 2.64, $p < .001$]. Visual inspection of the grand average difference waveform peaked around 400 ms and showed a maximum deflection of about -2.2 μV at posterior midline site Oz.

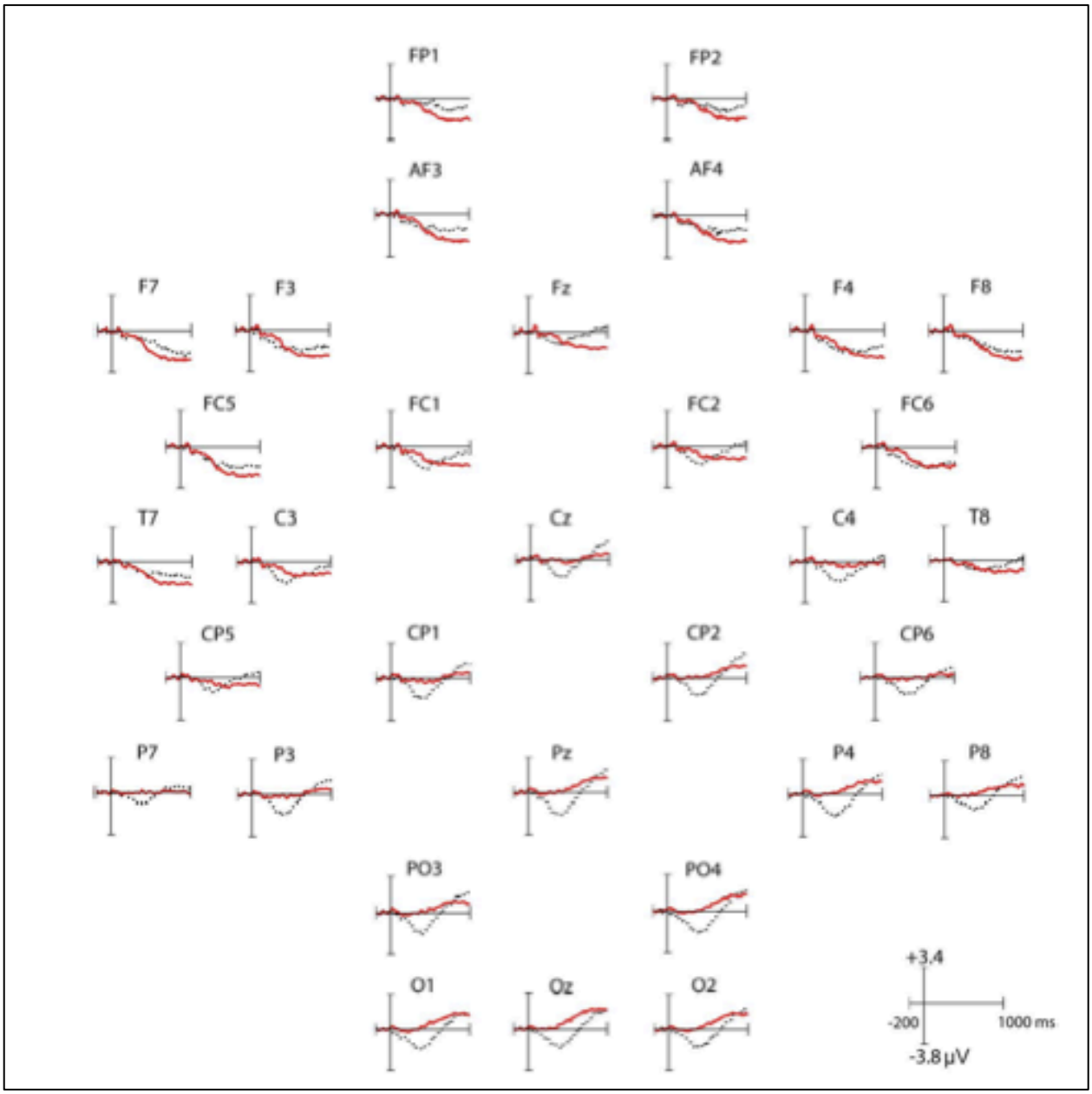

Figure 5.2 - Effect of cloze probability. Low cloze words (black dashed) show greater negativities than high cloze words (solid red) between 230-510 ms after word onset (negative plotted down).

Variant Effect 550-730ms (word onset)

After the main effect of Cloze ended around 510 ms, we identified a window between 550-730 ms after word onset which showed a significant main effect of Variant. As illustrated in Fig. 5.3,

IN' words show greater mean negativities than ING words [-.713 μV difference, F(1,27) = 6.92, MSE = 16.46, p = .014]. Visual inspection showed the Variant effect was largest at anterior left electrode sites, showing a maximum negative peak deflection at 680 ms after word onset of about -1.3 μV.

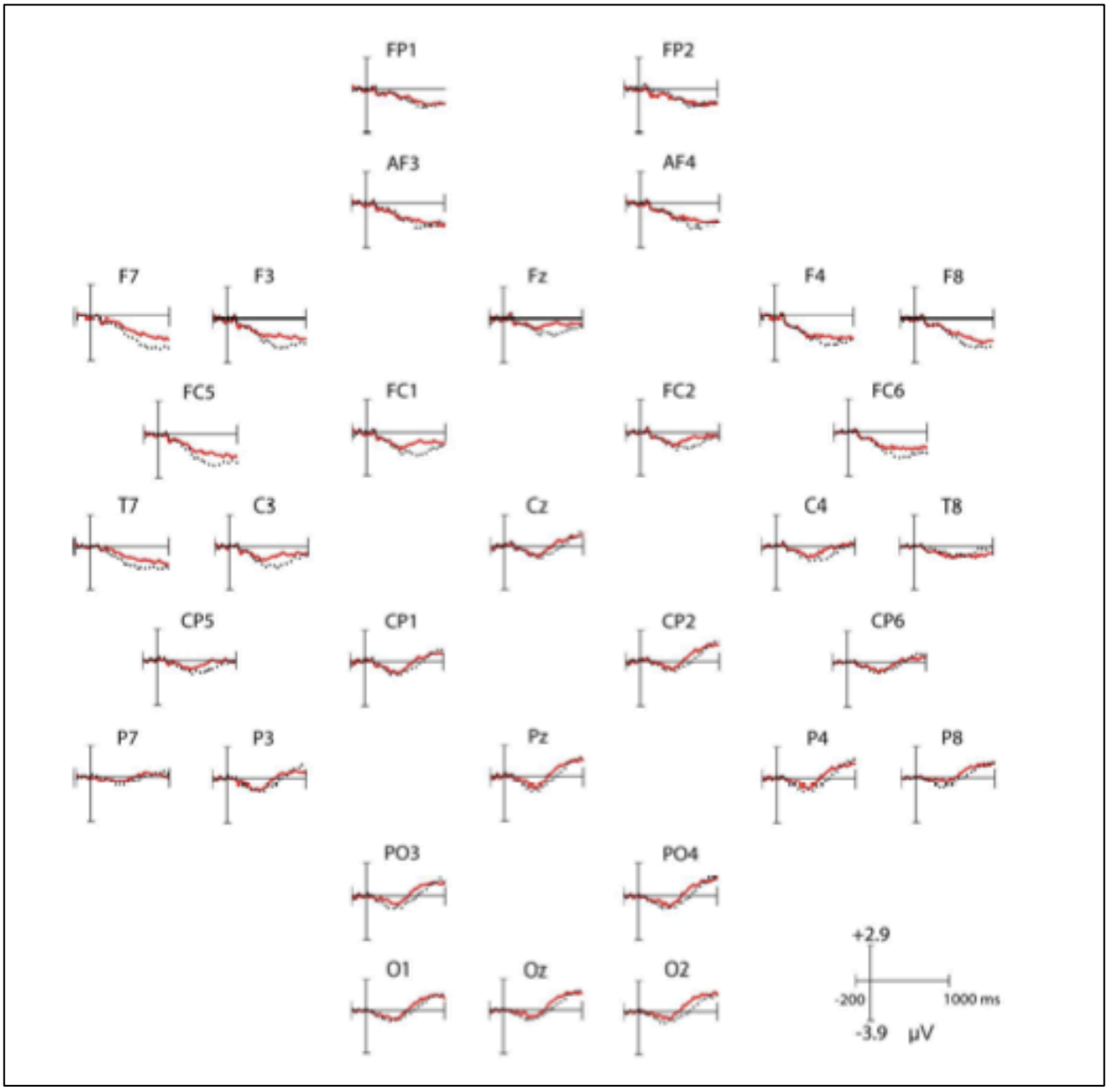

Figure 5.3 - Variant effect at word onset. IN' words (black dashed) are more negative than ING words (solid red) between 550-730 ms after word onset (negative plotted down).

Variant Effect 170-440 ms (morpheme onset)

In order to provide a more precise characterization of the timing of these variant effects, we conducted an additional analysis with the ERPs time-locked to the acoustic onset of the ING/IN' allomorph. Visual inspection and subsequent statistical analysis of these data identified a window between 170-440 ms, in which we observed a main effect of Variant. As illustrated in Fig. 5.4, IN' words (dashed black) are more negative than ING words (solid red) [-.794 μV difference, $F(1,27) = 16.05$, MSE = 8.81, $p < .001$]. As expected from the word onset analysis, visual inspection showed the Variant effect was largest at anterior left sites, showing a maximum negative peak deflection of about 2.0 μV at about 300 ms after morpheme onset.

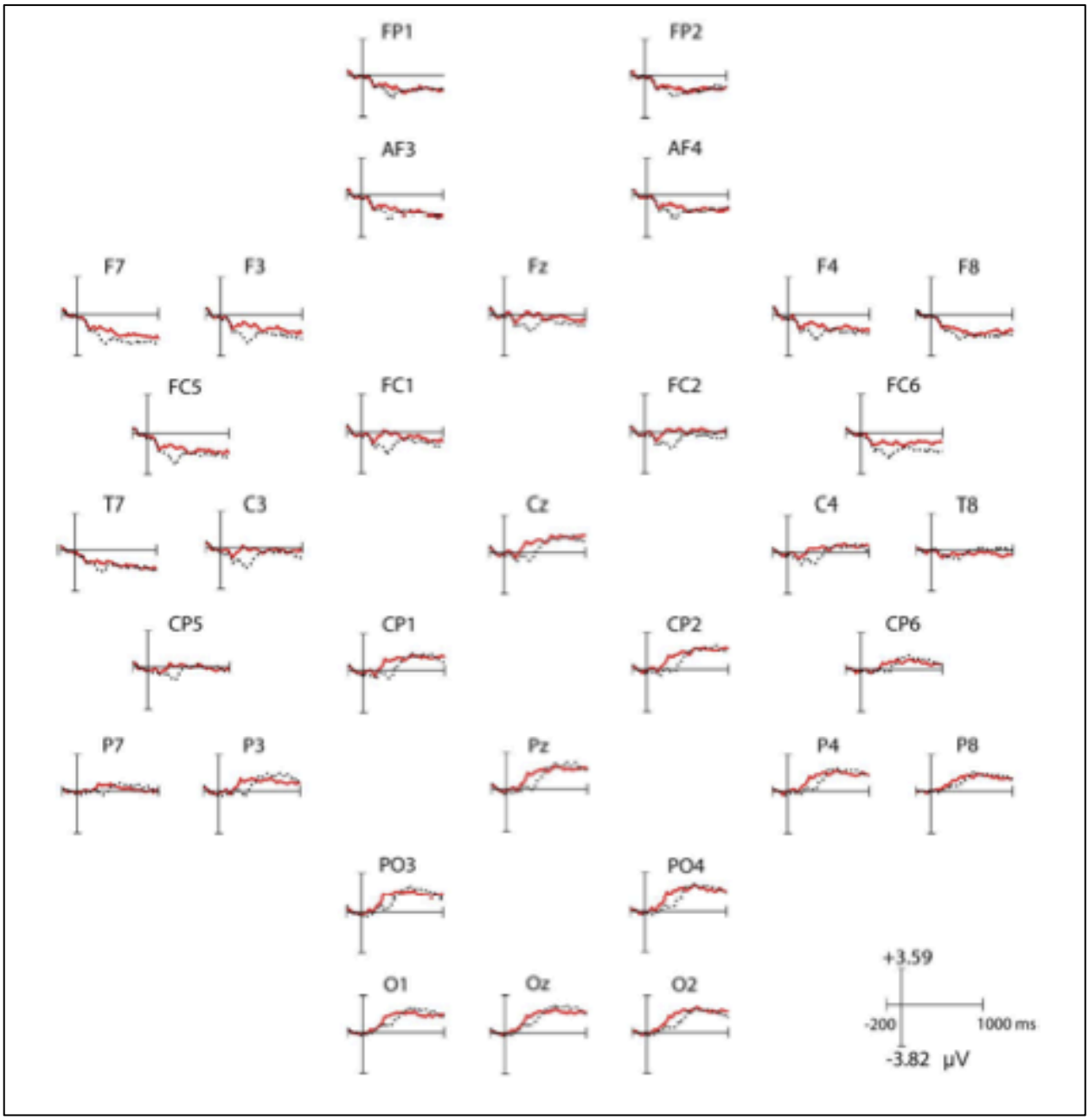

Figure 5.4 - Variant effect at ING/IN' onset. IN' words (black dashed) are more negative than ING words (solid red) between 170-440 ms after morpheme onset (negative plotted down).

Variant x Congruency 160-350 ms (morpheme onset)

With data locked to the onset of the ING/IN' morpheme, we identified a window between 160-350ms with a significant interaction of Variant x Congruency x Cloze x AP [F(1,27) =

9.276, MSE = 0.688, p = .005]. Pairwise comparison showed Informal→IN' was more negative than Formal→ING at posterior sites (0.70μV difference, p=0.032) for high cloze words and at anterior (1.58μV difference, p<0.001) and posterior sites (0.92μV difference, p=0.035) for low cloze words. Formal→IN' was more negative than Informal→ING at anterior sites (0.94μV difference, p=0.042) for high cloze words. Informal→IN' was more negative than Formal→IN' at anterior sites (0.87μV difference, p=0.015) for low cloze words. No significant differences were observed between Formal→ING and Informal→ING. The ordering of register effects by experimental condition are illustrated in Table 5.2.

Table 5.2 - Ordering of conditions by mean waveform amplitude. Rows display largest observed difference between that condition and the condition on the preceding row. For ease of interpretation, cloze probability is not displayed.

	Condition	μV	Difference
A.	**Formal→ING**		
B.	**Informal→ING**	-.30 (n.s.)	Row B – Row A
C.	**Formal→IN'**	-.94 (p=.042)	Row C – Row B
D.	**Informal→IN'**	-.87 (p=.015)	Row D – Row C

5.5 Discussion

In this study, we crossed semantic and sociolinguistic expectancies by manipulating scripted spoken passages with sentence-final critical ING/IN' words. We observed three primary effects: (1) an effect of cloze probability between 230-510 ms, with low cloze words showing greater posterior negativities compared to high cloze words; (2) an effect of sociolinguistic variant, with vernacular IN' variants eliciting greater negativities than canonical ING variants. This effect was observed between 550-730 ms with ERPs time-locked to word-onset and between 170-440 ms with data locked to morpheme onset; and (3) an interaction of Variant x Congruency x Cloze x

AP between 160-350 ms showing a hierarchy of N400-like negativities by condition, with Formal→ING and Informal→ING being the most positive, followed by Formal→IN', which in turn was followed by Informal→IN', the most negative condition. In the following discussion, we consider each of these effects before turning to the general discussion.

Cloze Probability and Sociolinguistic Variant Effects

When our analyses were based on ERPs time-locked to word onset, we found an N400-like effect for cloze probability, with low cloze words showing increased posterior negativities compared to high cloze words (Fig. 5.2). The distribution and amplitude of the cloze effect was comparable to those reported in other ERP studies of cloze probability (Delong et al., 2005; Kutas & Hillyard, 1984) and its topographic distribution was similar to those reported for semantic anomalies encountered in discourse and sentence-level processing (Hagoort et al., 2004; Van Berkum et al., 2003). As will be discussed shortly, posterior negativities in the N400 time window, such as those elicited by semantic anomaly and low cloze probability are often interpreted as indices of language processing difficulties or increased cognitive effort in semantic conceptual integration (Brown & Hagoort, 1993; Osterhout & Holcomb, 1992).

Similar to the N400 effect of cloze probability, we also observed an effect of sociolinguistic variant, with vernacular IN' variants eliciting a broadly distributed negativity compared to canonical ING variants. With ERPs locked to word onset, the effect of variant (550-730 ms, Fig. 5.3) occurred after the effect of cloze probability (230-510 ms, Fig. 5.2). At first glance the timing of the cloze probability and sociolinguistic variant effects might suggest that the sociolinguistic aspects of speech are processed after semantic processing, thus providing support for a two-step model of language processing. However, we believe that this is not the

actual case, but rather a reflection of the inherent properties of our spoken language stimuli. In English, the progressive morpheme is a suffix that attaches to the end of the verbal stem (e.g., *rain* + ING/IN', *storm* + ING/IN'). As such, any observed effect of variant would be expected to only occur after initiation of verb stem processing – when the acoustic onset of the ING/IN' allomorph is actually encountered in the temporal speech stream. Our gating study, for example, established that listeners could only reliably identify whether a word ended in ING or IN' approximately 360 ms after the acoustic onset of the verb stem. A more precise measure of the timing of variant effects, we reasoned, could be achieved by time-locking the ERPs to the onset of the actual ING/IN' allomorph instead of word onset.

Given the temporal variability between word and allopmorph onset, when the ERPs were more precisely time-locked to the ING/IN' morpheme itself, the effect of variant emerged earlier than when locked to the acoustic onset of the verb stem (170-440 ms, Fig. 5.4 vs. 550-730 ms, Fig. 5.3). As illustrated in Fig. 5.5, there are several key similarities and differences between the cloze effect and the variant effect. Both effects emerged quickly after the acoustic onset of the word (230 ms) or variant (170 ms). Both showed main effects of similar duration (Cloze, 280 ms; Variant 270 ms). Both effects showed a grand average peak difference of about 2 μV, but mean negativities were larger for the cloze effect (-1.341 μV) compared to the variant effect (-.794 μV). Whereas the variant effect peaks around 300 ms, the cloze effect peaks around 400 ms. In contrast to the broad similarities in wave morphology, amplitude, and timing, the difference in distribution is striking. Whereas the N400 cloze effect is largest at posterior sites, the variant effect has a more anterior left focus, suggesting at least partially distinct neural regions are recruited while processing the sociolinguistic aspects of speech.

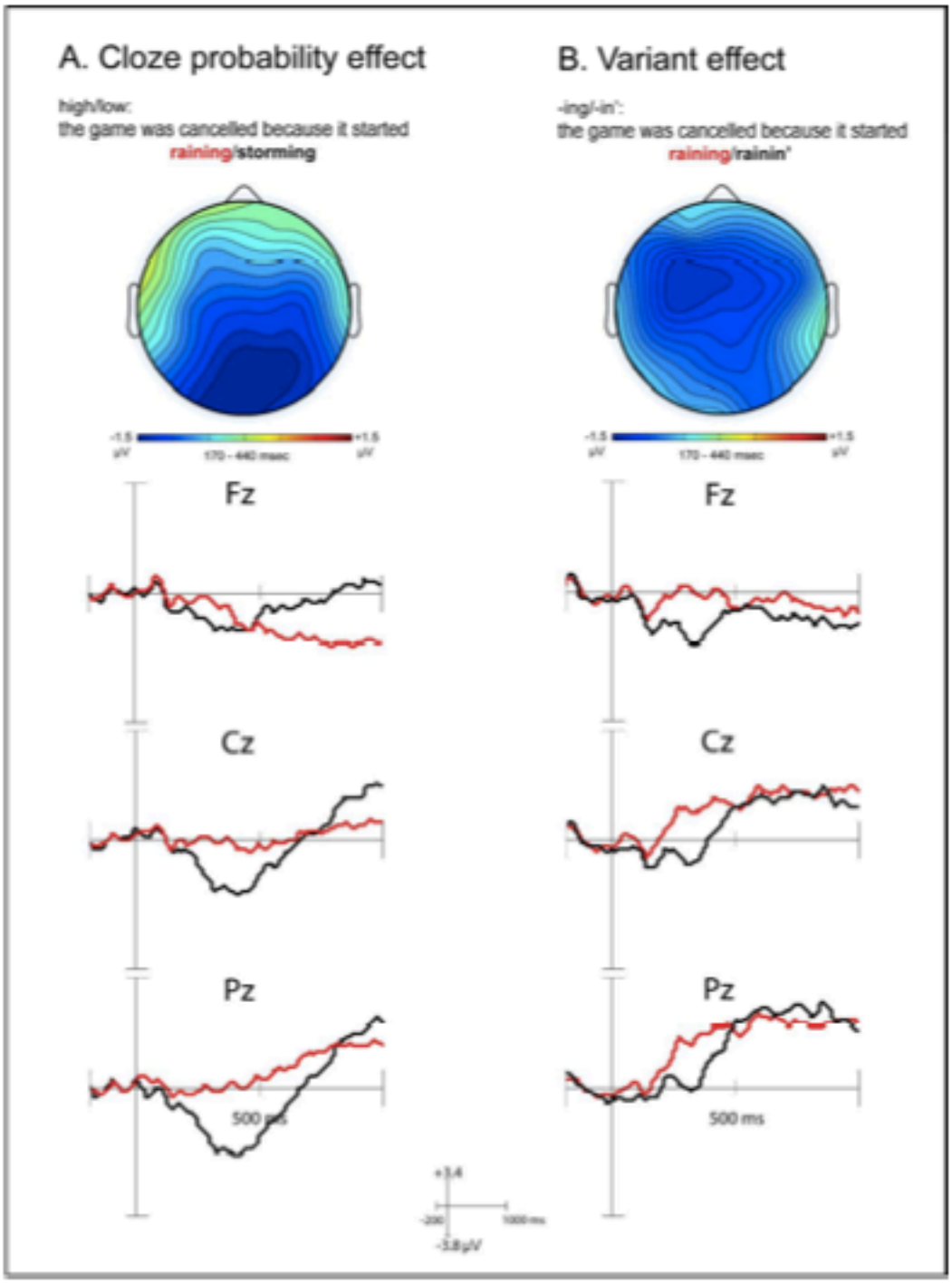

Figure 5.5 - Comparison of cloze effect and variant effect. Column A shows close-ups of high cloze (solid red) and low cloze (solid black) words at midline sites Fz, Cz, and Pz. The scalp plot shows the mean difference (low cloze – high cloze) between 170 – 440 ms after word onset. Column B shows close-ups of ING words (solid red) and IN' words (solid black) at midline sites Fz, Cz, and Pz. The scalp plot shows the mean difference (IN' – ING) between 170-440 ms after morpheme onset (negative plotted down).

While there are disagreements over the actual mechanisms involved, there is a consensus that centro-parietal negativities within this time-window reflect processing difficulties or increased cognitive effort that results from encountering unexpected or semantically anomalous words (Brown & Hagoort, 1993; Osterhout & Holcomb, 1992). One favored interpretation is that the N400 is an index of combinatorial processes involved in conceptual semantic integration, with waveform amplitude being a goodness-of-fit measure of word-meaning to larger conceptual representations (Brown & Hagoort, 1993; Hagoort & Van Berkum, 2007; Van Berkum et al., 1999). A contrasting view of the N400 argues that it indexes lexical access proper as discussed by Kutas and Federemeier (2000) and Lau et al. (2008). In this interpretation, expectations built up during online processing spread activation to features in long-term memory which facilitates lexical access when expected words are encountered. That is to say, the previously encountered context "allows pre-activation of relevant lexical or conceptual features, making lexical access less effortful" (Lau et al., 2008).

One well-studied factor that facilitates lexical access and which reduces N400 amplitude is lexical frequency. A number of studies, for example, have shown that high frequency words show an attenuated N400 compared to low frequency words (e.g., Barber, Vergara, & Carreiras, 2004; Smith & Halgren, 1987). In this "lexical account" then, canonical forms such ING may be more strongly represented than vernacular variants because they are more frequently encountered and produced by our college-educated participants, and thus the less-frequent vernacular variants might be expected to require more effort to access. Indeed, some support for this position was found in the two production studies we conducted with age-matched subjects. Sociolinguistic interviews conducted between close friends on casual conversational topics, revealed vernacular IN' frequency at approximately 10% of the total number of tokens. Similarly, the video

description task, which required participants to narrate a cartoon video in real-time, revealed a similarly small frequency of vernacular tokens. Though both tasks were designed to elicit informal, relative unmonitored vernacular speech, it is somewhat surprising to note the relative lack of IN' variants in this community which contrasts with reports of other speech communities (e.g., Abramowicz, 2007; Hazen, 2008; Wolfram & Christian, 1976).

An alternative position argues that N400 amplitude reflects the relative ease or difficulty of integrating a word's meaning into higher order conceptual structures. Although variants such as *working* and *workin'* are synonymous (i.e. they mean the same thing), they index different speech registers (among other factors) and thus have the potential to convey different socially indexical information (Labov et al., 2006; Wald & Shopen, 1985). One plausible account of the vernacular variant effect is that listeners construct default formality expectations cued by larger contextual factors such as the setting and audience. In this case, when variants that typify informal vernacular speech are encountered in a relatively formal setting, integration difficulties may arise due to mismatches with social expectations that reside as part of higher-order conceptual representations. It is important to note, however, that these sociolinguistic expectations may be heightened in this particular case by the context of the "psychology experiment" itself – a setting which has its own rather formal speech registers, that favor standard variants and clear articulation. Thus by this "social-expectancy" account, the processing difficulties arise because of a conflict between the predicted level of formality elicited by the broader context (largely construed) and the word form which is actually encountered in that context.

An alternative explanation of these difficulties in integration processes may be due to perceptibility issues of the acoustic properties of these spoken language stimuli. As discussed in

the methods, we observed non-significant trends in mean F0 (a measure of prosody) and intensity between ING and IN' critical words. While not statistically significant at the $p < .05$ level, the trends do indicate that these factors may be, in part responsible for the pattern of data observed. However, as discussed in the following two sections, higher order interactions of variant, congruency, and gender alternatively suggest these data reflect the processing difficulties of the sociolinguistic aspects of speech.

Congruency & Variant Interaction Effects

In addition to the main effect of variant, we also observed a hierarchy of N400-like negativities by condition, with Formal→ING and Informal→ING being the most positive, followed by Formal→IN' and finally by Informal→IN', the condition which elicited the greatest negativity. As illustrated in Table 5.3, the observed hierarchy of results stands in partial contrast to the ordering of effects predicted by both the congruency and the frequency-based accounts of sociolinguistic processing. As will be recalled, the congruency account predicts increased processing costs for critical words that "violate" or are incongruent with the preceding register context compared to variants that match the register context. Although this account correctly predicts eased processing for congruent variants in formal register contexts (Formal→ING), the congruency account further predicts relatively eased processing for vernacular IN' variants in informal contexts (Informal→IN'), a prediction which is not borne out by these data.

Table 5.3 - Ordering of predicted and observed effects by waveform amplitude

μV	Congruency Account	Frequency Account	Observed Hierarchy
+	Formal→ING	Formal→ING	Formal→ING = Informal→ING
	Informal→IN'	Formal→IN'	Formal→IN'
	Informal→ING	Informal→ING	

-	Formal→IN'	Informal→IN'	Informal→IN'

In contrast to the congruency account, the frequency-based account predicts processing cost as a function of overall frequency of vernacular IN' variants in each of the passages. As seen in Table 5.3, the frequency account largely predicts the ordering of effects observed in our data, with negative waveform amplitude increasing from Formal→ING to Formal→IN' to Informal→IN' conditions. The frequency account failed, however, to correctly predict canonical ING variants in informal contexts, which showed no significant difference from canonical ING variants in formal contexts (i.e. Formal→ING = Informal→ING). We speculate that the differences observed in Labov, et al. (2011 who reported attitudinal judgment as a function of vernacular IN' frequency and the present study may be due to the types of measures used. Whereas Labov and colleagues used an offline matched guise task that may encourage more deliberative and evaluative language processing, our study relied on direct measures of brain activity during real time language processing. The question that emerges from our data, is why would informal register contexts increase N400 negativities on vernacular IN' critical words compared to formal contexts, but not on canonical ING words?

We propose an alternative accounting of these data. If canonical ING is the default variant of this speech community, listening to passages characterized by this variant should minimize language processing costs. Thus, while listening to formal register contexts, listeners' sociolinguistic expectations are being met while the discourse context continues shaping expectations for highly predicted critical words and default ING variants. When less predicted critical words and vernacular variants are encountered in this formal register context, a processing cost is incurred resulting in heightened N400-like potentials compared to the expected forms. In contrast, consider listening to an informal register context characterized be a series of

IN' variants. Because of the relative rarity of the vernacular variant in this speech community, hearing these informal register contexts may serve to heighten attentional processes, increasing expectations for the canonical ING variant above default levels. When these sociolinguistic expectation are met (Informal→ING) processing proceeds as usual; However, if a vernacular IN' critical variant is encountered (Informal→IN') in this state of heightened expectations for the canonical ING form, language processing appears to become more taxing as indicated by increased N400-like negativities compared to encountering the same vernacular IN' variant in a formal register context (Formal→IN').

General Discussion

According to our results, the brain begins to integrate word meaning and indexical meaning within 170-210 ms after encountering the acoustic onset of the critical word or variant. Our study demonstrates that while listening to speech for comprehension, brain processes are sensitive to sociolinguistic variants that can serve as cues to social information. The temporal and morphological characteristics of the brain response suggest that sociolinguistic variation is rapidly brought to bear in the on-going sentence interpretation and is treated akin to word-level semantics, suggesting there may be few differences between formal semantic meaning and social meaning conveyed by how a word is pronounced. This is novel, because in contrast to word meaning, any social information indexed by variation is only probabilistic in nature (Labov, 1966). That is to say, informal social contexts, such as conversations among peers, only tend to elicit more vernacular variants, while formal contexts only tend to elicit more canonical variants.

In contrast to those who have suggested that the interpretation of pragmatic and contextual properties of language are secondary to the computation of local sentence meaning

(Lattner & Friederici, 2003; Regel, Gunter, & Friederici), our results join a growing body of evidence that support a single-step integration process that simultaneously takes into account both local and global factors during language processing (Hagoort et al., 2004; Hagoort & Van Berkum, 2007; Ozyurek, Willems, Kita, & Hagoort, 2007; Van Berkum et al., 2008; Van Berkum et al., 2003). Our results extend these findings by showing that listeners' brains track sociolinguistic cues in real-time in the service of language comprehension and that when stigmatized vernacular variants are encountered, language processing is taxed.

The partially different distributions of the cloze probability and variant effects indicate that partially distinct brain systems are involved (Osterhout & Hagoort, 1999). We speculate on three possible reasons why processing the sociolinguistic aspects of speech recruits partially different neural systems. While listening to speech in real time, the effect of cloze probability is elicited when the first phonological segments of a word are encountered that mismatch those that had been predicted (e.g., when [s]torming is encountered, when [r]aining was predicted). The resultant N400 effect typically becomes significant around 200 ms after this initial acoustic mismatch and returns to baseline around 500 ms. However, by the time the critical ING/IN' morpheme is encountered and recognized (~360 ms from word onset), the cloze effect is already ending and processes involved in sentence wrap-up may already be well underway. The different distribution of the variant effect may be partially attributed to this additional processing. Second, processes involved in morpheme decomposition may be involved during variant processing. A study by Leinonen et al. (2009), for example, reported an N400 response to auditorily presented morphologically inflected words with a more anterior distribution similar to the one observed in our study. Finally, the differences in distribution may reflect representational difference in

semantic meaning and socially indexical meaning that recruit different brain regions, particularly so for women.

The core finding, that vernacular word variants tax language processing, raises several important issues which will need to be addressed in future research. If participants pre-activate default sociolinguistic expectations, which are then "violated" when vernacular tokens are encountered, how then do other social and linguistic contexts modulate this response? For example, the stimuli used in our experiment, were all short, semantically rich vignettes about everyday people doing everyday things (Appendix A). Would more formal genres (e.g., political speeches) elicit a heightened effect of variant, similar to the negative reactions elicited by Sarah Palin's speech during the 2008 vice presidential debates (Purnell et al., 2009)? In contrast, would more naturalistic discourse reduce N400 negativities for vernacular variants? For example, the stimuli used in our experiment were not snippets of spontaneous, extempore speech, but rather speech read from scripts. As others have noted, spontaneous conversational speech compared to read speech is characterized by more frequent use of vernacular variants (e.g., Labov, 1966). As such, the use of these forms in our read speech stimuli may serve to increase sociolinguistic expectancies above levels found in natural conversational discourse contexts.

Finally, additional work is needed to explore the degree of variant modulation for social groups that differ in their frequency of variant use. For example, speakers of other regional dialects of English (e.g., southern and Appalachian dialects) use vernacular IN' variants much frequently (Hazen, 2008; Houston, 1985) than speakers of standard dialects. Similarly, holding dialect region constant, men compared to women are more likely to use vernacular IN' rather than canonical ING variants in their own speech. As such, speaker/listeners with these production characteristics might not be as sensitive to hearing the stigmatized variant as it occurs more often

in their own speech. In addition, future work will need to examine whether naïve listeners taken into account the speaker of the dialect uttering the variant during real-time language processing. For example, is hearing a Southern English speaker utter a vernacular IN' form less taxing on language processes than hearing a speaker of a standard dialect uttering the same variant?

5.6 Summary

In order to understand the speech of others, the meanings of individual words must be integrated into higher-level conceptual structures. When unexpected words are encountered, additional cognitive resources are recruited in this endeavour. Our study shows that these same processes operate on variable word forms (*working/workin'*) that differentially index social information. This study provides strong evidence that the cognitive mechanisms that support language comprehension are sensitive not just to *what* is said, but also to *how* it is said. Results from this study found that brain processes are sensitive to sociolinguistic variants that can serve as cues to social information as early as 170 ms after the acoustic onset of the variant. The early onset, polarity, and distribution of the effects suggests that sociolinguistic variation is rapidly brought to bear during sentence interpretation and is processed similarly to word-level semantics. This provides support for a single-step model of language comprehension in which integration process simultaneously take into account both semantic and social meaning during language processing. Importantly, however, these integration processes are further constrained by the preceding register context uttered by the speaker, with informal contexts increasing N400 negativities for vernacular variants compared to formal register contexts.

6. Speaker Dialect and the Perception of Sociolinguistic Variation

6.1 Introduction

In the study of human communication, psycholinguistic models of language processing have primarily focused on word and sentence-level processing. When issues of pragmatics and social context have been considered, they have traditionally been argued to occur after local sentence meaning has been computed (Cutler & Clifton, 1999; Lattner & Friederici, 2003). An increasing body of research, however, suggests that larger contextual factors, such as world-knowledge and speaker-knowledge, are integrated into the unfolding sentence interpretation by the same cognitive processes that support local lexical-semantic integration (Hagoort, Hald, Bastiaanesen, & Petersson, 2004; Van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008) (For review, see Chapter 4). What remains relatively unknown and underexplored in these accounts, is when and how social information cued by sociolinguistic variants and speaker dialect affects language processing. The experiment reported in the present chapter measures changes in brain activity using electrophysiological scalp recordings to assess the real-time integration of sentences spoken by different dialect speakers that differ in the realization of the English progressive inflectional morpheme (ING).

During language processing, the meanings of individual words are combined with morphosyntactic information in order to support comprehension. However, in addition to information conveyed through syntax and semantics, *how* something is said can convey

additional social and contextual information. For example, an employee, when asked about his weekend, might report to his boss: "*Last weekend we went hunting.*" However, the same speaker talking among peers might say: "*Las´ weeken´ we went huntin´."* One way in which these speech acts differ, is in the realization of specific sociolinguistic variables or "alternative ways of saying the same thing"(Fasold, 1990) (e.g., pronouncing the word *hunting* with its canonical velar ING form or using its vernacular apical IN' pronunciation *huntin'*). Similarly, in terms of regional dialectal variation, one might expect a speaker from the Southern United States to be more likely pronounce the word *huntin'* whereas a speaker from California speaker might be expected to use the canonical pronunciation *hunting*.

One of the primary findings in sociolinguistics is that variation follows probabilistically regular patterns (Guy, 2007). How a specific sociolinguistic variable is realized in a particular speech context is probabilistically constrained by linguistic factors, such as lexical class and phonological environment, and external socio-stylistic factors, such as the speaker's gender and dialect, as well as the relative formality of the register (e.g. a speech read in front of a formal audience versus a casual conversation among friends). In the decades following Labov (1966), the variationist paradigm has witnessed an impressive accumulation of knowledge characterizing influences on the *production* of linguistic variation. Relatively lacking from this body of research, however, is an account of the cognitive mechanisms that support real-time language processing of socially indexical word variants and dialectal variation. The present experiment addresses this gap by using event-related brain potentials to examine the processing costs associated with variable word forms spoken by speakers of different dialects.

In a previous study (Chapter 5), we recorded participants' ERPs while they listened to short passages that varied in (ING) realization and cloze probability. In these spoken passages,

some verb forms were spoken with the canonical velar ING pronunciation and others were realized with vernacular alveolar IN'. We reported three primary findings: (1) a classic N400 for low cloze probability words compared to high cloze probability words; (2) a widely distributed N400-like negativity for vernacular IN' variants compared to ING variants; and (3) heightened N400 negativities for vernacular IN' variants when preceded by an informal register context compared to a formal context. We interpreted these findings as support for a single-step model of language processing, one that rapidly takes into account the pragmatic and social aspects of language use during real-time language comprehension. Moreover, the observed interaction of register and variant further suggested that local linguistic contexts can in some cases modulate expectations for variable word forms.

One factor that the previous study fails to address is whether local linguistic contexts alone are sufficient to entirely override broader sociolinguistic expectations for canonical variants. For example, it could be argued that the vernacular variant effect observed in the previous experiment is attributable to expectations engendered by the psycholinguistic experiment itself[3] – a rather formal social context which may be increasing expectations for formal word variants over vernacular forms. In an attempt to broaden the effects observed in ERP experiment #1, in this second ERP experiment, we sought to elicit an N400 like effect not just for vernacular IN' forms as in our previous work, but also for canonical ING forms as well.

To this end, we attempted to replicate and extend our previous work by examining how speaker dialect modulates the electrophysiological brain response to sociolinguistic variation. Specifically, we sought to address whether the brain uses information conveyed by speaker dialect in order to predict which variable word form is a speaker is likely to utter during real-time language processing. It has been well established in the production literature that (ING)

[3] I thank an anonymous reviewer of this study for bring my attention to this issue.

realization is constrained by speaker dialect. In the United States, speakers of Southern American English (SAE) tend to use the vernacular IN' variant at much higher frequencies than speakers of other U.S. regional dialects (Campbell-Kibler, 2007; Mencken, 1963; Wolfram & Christian, 1976). Labov (2001), for example, has claimed that in the Southern United States, vernacular IN' "is used *almost exclusively* in speech, even in the most formal kind" (emphasis added, p. 90). The question we are addressing here is whether speaker dialect engenders expectations for variable word forms. That is to say, do naïve listeners "know" that Southern speakers are more likely to use vernacular IN' variants compared to speakers of more standard dialects? Does the listener's brain make use of these probabilistic dialectal tendencies to help predict which variant a speaker is likely to utter? From a theoretical perspective, such a mechanism would obviously play an important role in real-time language processing by helping constrain the lack of invariance between variable acoustic realizations and underlying abstract form.

The current experiment attempts to answer these questions, by examining whether speaker dialect engenders listener expectations for dialect typical (ING) variants. Specifically, we ask if listening to a Californian speaker increases expectations for ING words? Likewise, does listening to a Southern dialect speaker increase expectations for IN' forms? As with our previous study, subjects listened for comprehension to short, semantically rich stories that varied in (ING) realization. Half the critical words were realized with ING and half realized as IN'. All sentence final words were highly predictable, as established by a previous cloze probability task (mean cloze probability 74.9%, range 20.7% - 100%). As discussed in Chapter 5, all critical words were either preceded by a formal register context containing four embedded ING words or by an informal context containing four IN' words. In contrast to ERP experiment #1, however, in the present experiment, half of the passages were read by Californian speakers and half were read by

Southern speakers. In summary, we used a 2x2x2 factorial design: Variant (ING/IN') x Register (Formal/Informal) x Speaker Dialect (Californian/Southern).

6.2 Predictions

Based on the previously discussed sociolinguistic and ERP literature, as well as the results from our first ERP experiment (Chapter 5), we predicted the following: (1) increased N400-like negativities for vernacular IN' variants and informal contexts compared to ING variants and formal contexts when spoken by California dialect speakers; (2) increased N400 negativities for canonical ING variants and formal contexts compared to IN' variants and informal contexts when spoken by Southern dialect speakers. In addition, we predicted a hierarchy of N400-like potentials as a function of speaker dialect, variant, and preceding register context. We predicted N400 amplitudes to increase as conditions diverged from the most dialect typical to the least typical. As illustrated in Table 6.1, listening to a Californian speaker utter an IN' variant should by more taxing than an ING variant. Similarly, for register context, listening to a Californian speak an informal context should tax language processing more than a formal context. In contrast, listening to a Southern speaker use an ING variant should tax processing more than an IN' variant, and formal register contexts should be more taxing than informal contexts. As observed in ERP experiment #1, we further expected the critical variant compared to the preceding register context to have more of an effect on N400 amplitude, accounting for the observed hierarchy presented in Table 6.1.

Table 6.1 - ERP Predictions for Californian versus Southern dialect speakers, Dialect knowledge account.

	Californian Dialect	**Southern Dialect**	**N400 Amp.**	**Interpretation**

Dialect Typical	Formal→ING	Informal→IN'	Smallest	Least Taxing
	Informal →ING	Formal→IN'		
	Formal→IN'	Informal →ING		
Dialect Atypical	Informal→IN'	Formal→ING	Largest	Most Taxing

We also considered the possibility that results might pattern by listener familiarity to speaker dialect instead dialect typicality per se. As all subjects were Californian natives, this account assumes greater familiarity and exposure to Californian speakers compared to Southern speakers. In general, this account predicts heightened N400-like potentials for speakers of non-familiar dialects (i.e. SAE) compared to speakers of familiar dialects (i.e. Northern Californian English). Moreover, although this account predicts a hierarchy of N400 potentials by condition for Californian speakers (as illustrated by the pattern in ERP experiment #1), unfamiliar Southern speakers might pattern differently. If California listeners have few sociolinguistic expectations for unfamiliar dialect speakers, one might expect no differences in N400 amplitude by condition for Southern speakers. Alternatively, listeners might rely on their own dialect-specific sociolinguistic expectations when attending to the speech of unfamiliar speakers, which predicts similar hierarchies for both Californian and Southern speakers, but with Southern speakers eliciting greater negativities than Californians for each condition (Table 6.2).

Table 6.2 – ERP Predictions for Californian vs. Southern Speakers, Dialect familiarity account

	Californian Dialect	**Southern Dialect**	**N400 Amplitude**	**Interpretation**
Dialect Typical	Formal→ING	Formal→ING	Smallest	Least Taxing
	Informal→ING	Informal→ING		
	Formal→IN'	Formal→IN'		
Dialect Atypical	Informal→IN'	Informal→IN'	Largest	Most Taxing
N400 Amp.	Smaller N400	Larger N400		
Interpretation	Less Taxing	More Taxing		

6.3 Methods

6.3.1 Stimuli

With the exception of the change in experimental design, we followed the same basic methodology detailed in Chapter 5. The 280 previously scripted passages were randomly assigned to six new speakers – three Californian speakers and three Southern speakers. All speakers were male and the two dialect groups were roughly balanced by age (range 25-30 years old). Californian speakers were all natives of Northern Californian, and the Southern speakers were natives of North Carolina who had moved to northern California within the past 2 to 3 years.

All speakers were recorded in a sound attenuated room with a professional microphone and Marantz PMD661 digital recorder. Speakers read from written scripts with line breaks to indicate pauses, and all speakers were asked to read with normal pitch and intonation. Each speaker read their passages in two different guises, using all ING words in one guise, and all IN' words in the alternative guise. Undergraduate research assistants in the Department of Linguistics assisted in digitally splicing the recordings to create recordings for four experimental conditions (i.e. Formal→ING, Formal→IN', Informal→ING, and Informal→IN'). Each passage was single .wav file that lasted between 10-25 seconds. Each stimulus had a trigger cue inserted at the onset of the ING/IN' nasal in the sentence-final critical word. Eight lists were created with 35 stimuli from each of the eight conditions. Lists were interspersed with 120 filler passages read

by the same six speakers. Presentation of stimuli was randomized and presented across 9 experimental blocks each approximately 10-15 minutes in length.

6.3.2 Participants

We recruited 38 new subjects to participate in this experiment. All subjects were undergraduate participants (female = 24) enrolled in Psychology and Linguistics courses at the University of California, Davis. Due to technical error, excessive artifacts, high alpha activity, and non-native English status, 19 participants were excluded from further analysis. As the length of the experiment was approximately 3.5 hours, including paperwork, EEG preparation, ERP recording, and two additional behavioral experiments, this was likely a contributing factor to the higher subject rejection rate of ERP experiment #2. The remaining 19 native-speaker participants (female = 11) ranged in age from 19 to 24 years and all were raised in Northern California (primarily the Bay Area and Central Valley).

6.3.3 EEG Recording and Analysis

Participants' electroencephalogram (EEG) was recorded from a 32 channel active electrode cap. Six external electrodes were placed on subjects for off-line signal processing and artifact detection: the left and right mastoid (reference electrodes), the left and right outer canthi (for saccades), and above and below the left eye (for blink detection). The EEG was re-referenced offline to the averaged left and right mastoids. Impedances were maintained below 20 k and EEG data was sampled at 256 Hz.

Subjects' EEG signal was filtered off-line with a band pass filter set to .01 and 30 Hz (Luck, 2005). The routines in the ERPLab plugin to EEGLab were used for artifact detection: (1) a moving 100 ms wide window peak-to-peak (voltage) threshold was set (100 μV) using 50 ms steps (erp_artmwppth); (2) Blink detection was calculated on the difference of the two VEOG channels, with the blink width set to 350 ms (erp_artblink); (3) Horizontal saccades were detected on the HEOG channel with a moving window of 200 ms with 50 ms steps with a threshold of 30 μV (erp_artstep). All trials were double checked by hand to ensure accurate artifact detection. Trials with artifacts due to eye blink, saccade, or movement were not included in analysis.

An average waveform was computed for each subject for each of the eight experimental conditions across four electrode regions: anterior left (AL) - Fp1, Af3, F7, F3, Fc1, Fc5; anterior right (AR) - Fc6, Fc2,F4,F8, Af4, Fp2; posterior left (PL) - Cp1, Cp5, P7, P3, Po3, O1; and posterior right (PR) - O2, Po4, P4, P8, Cp6, Cp2. Each epoch consisted of a 200 ms pre-stimulus baseline followed by a 1,000 ms post-onset window. Epochs were divided into 100 adjacent windows of 10 ms, and the mean amplitude for each window was entered into a repeated measure ANOVA using the SPSS statistics software package.

In order to fully analyze the factors and their interactions in this basic design – Variant (ING/IN') x Register (Formal/Informal) x Speaker Dialect (Californian/Southern) – it was necessary to conduct three separate ANOVAs in order to observe details of nested factors that are masked in any single analysis. Figure 6.1 illustrates the basic design of the stimuli – each stimulus read by a Californian or Southern speaker consisted of a short passage context containing four ING or IN' words, followed by a single sentence-final critical ING or IN' word. There exist three equally valid, statistically equivalent factorial designs that characterize these

stimuli: (1) Variant (ING/IN') x Congruency (Congruent/Incongruent) x Dialect (Californian/Southern); (2) Variant (ING/IN') x Register (Formal/Informal) x Dialect (Californian/Southern); and (3) Register (Formal/Informal) x Congruency (Congruent/Incongruent) x Dialect (Californian/Southern). Depending upon which design is used for statistical analysis, different significant interactions/main effects (but the with same total number of effects and with the same p values) can manifest because of the nested nature of this design. For example, a main effect of Variant under analysis #1 that shows greater negativities for IN' variants compared to ING variants, might emerge as a Register x Congruency interaction under analysis #3, with subsequent pairwise comparisons showing additional significant differences among conditions of nested variables masked by analysis #1. In order to facilitate presentation and discussion of findings, we present main effects and interactions using factorial design #1, described above. We supplement these findings by providing all the masked significant pair-wise comparisons that are driving these effects (designs #2 and #3 described above), with no correction for multiple pair-wise comparisons.

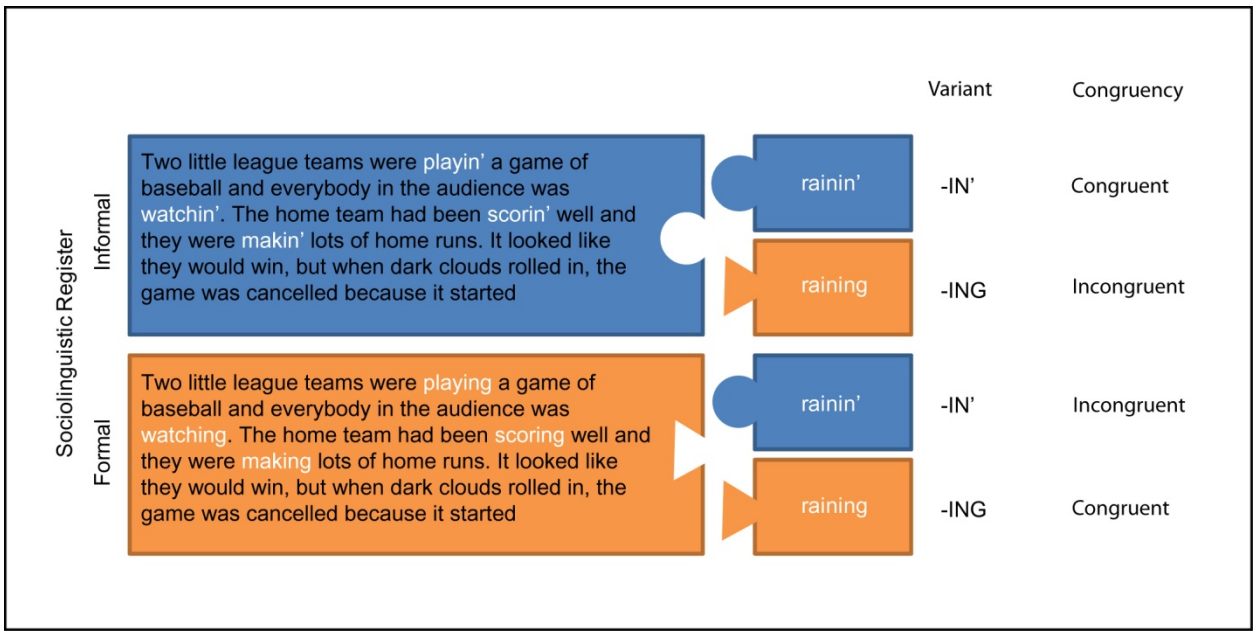

Figure 6.1 - Experimental stimulus design

6.4 Results

Variant Effect 260-510ms (morpheme onset)

We observed a main effect of Variant between 260 – 510 ms with a broad negativity peaking around 400 ms with an anterior left hemisphere maximum [Variant: $F(1,18) = 7.58$, MSE 30.48, $p = 0.013$]. As illustrated in Fig. 6.2 and Fig. 6.3, heightened negativities in this time window were observed for IN' variants compared to ING variants (mean difference 1.23 mV).

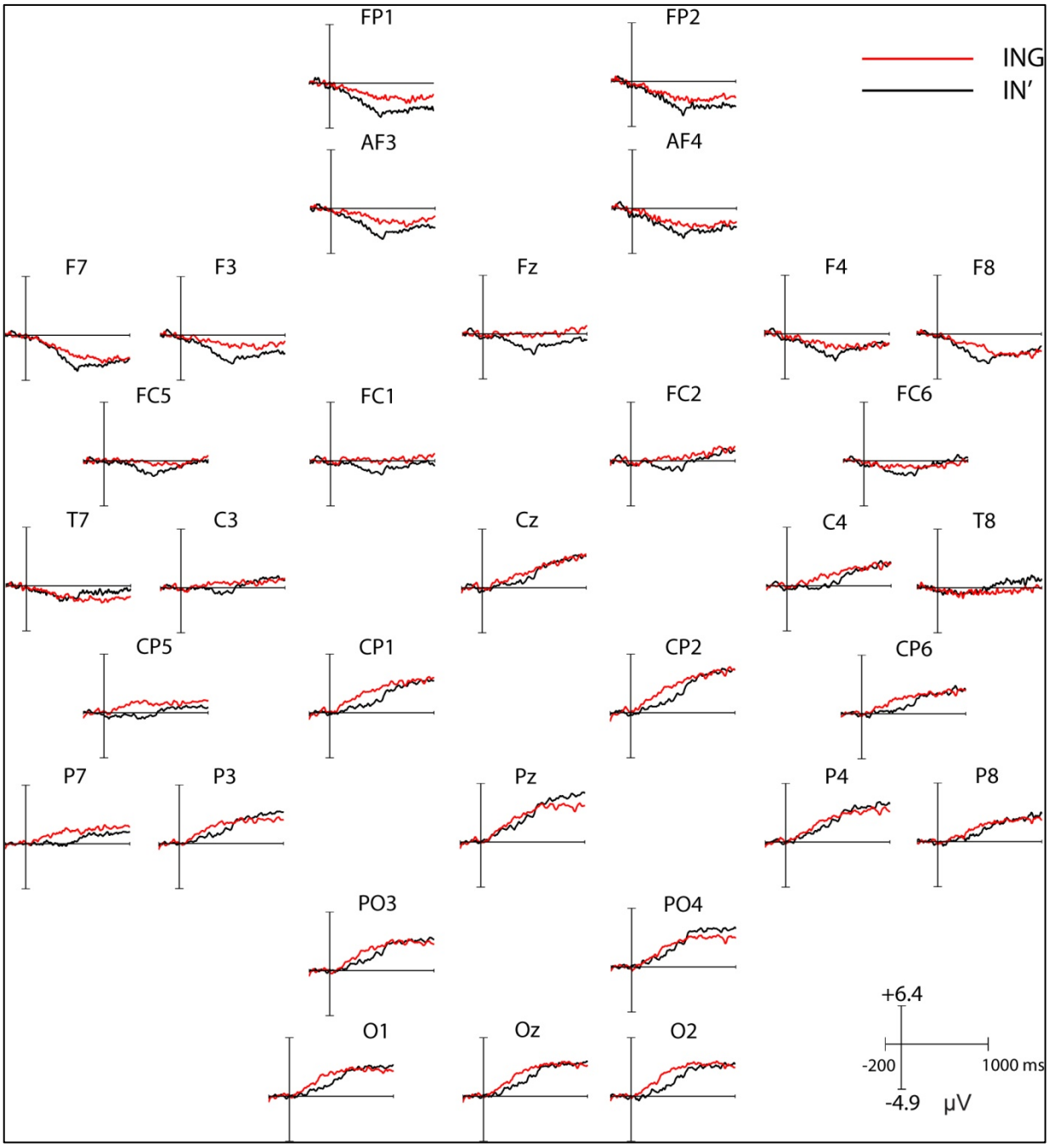

Figure 6.2 - Effect of sociolinguistic variant. Grand average ERP (n=19) for ING (red) and IN' critical words. Positive plotted up.

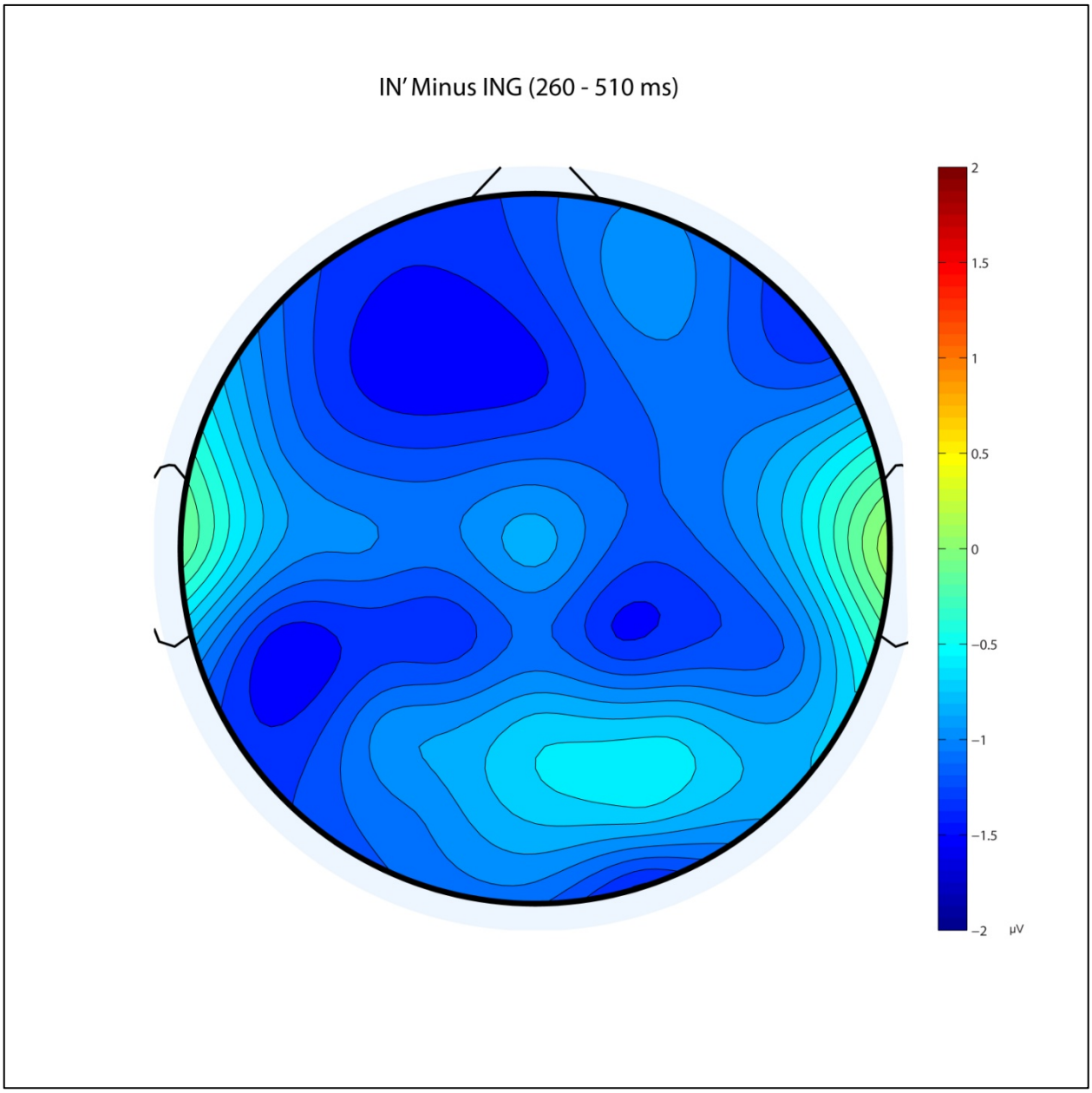

Figure 6.3 - Grand average difference (n=19), IN' minus ING words (260-510 ms).

Variant x Dialect Effect 250-370 ms (morpheme onset)

Importantly, we observed a significant Variant x Dialect interaction between 250 to 370 ms after ING/IN' acoustic onset [F(1,18) = 5.162, MSE 25.71, p = 0.036]. Pairwise comparisons showed increased negativities for IN' compared to ING variants spoken in a Californian dialect (p = 0.003). As illustrated in Fig. 6.4, IN' variants spoken by Californians (left) elicited mean negativities 2.21 mV greater than ING variants spoken by these same Californian speakers. In contrast, listening to Southern speakers (right), elicited no significant amplitude differences between IN' and ING variants. However, Within the 250-370 ms temporal window, we did observe a statistical trend of greater negativities for ING variants spoken by Southern speakers compared to California speakers (p = 0.063, mean difference 1.20 mV, Fig. 6.5).

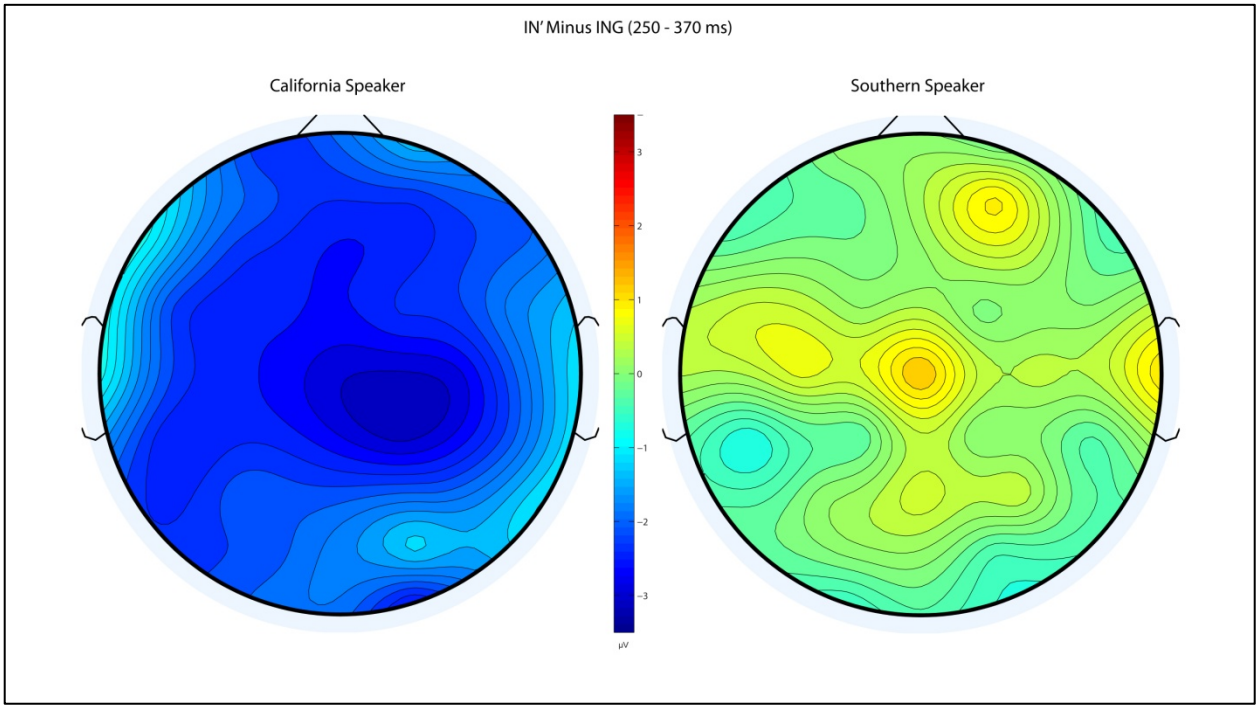

Figure 6.4 - Grand average ERP (n=19) difference, IN' minus ING (250-370 ms), for Californian speakers (left) and Southern speakers (right).

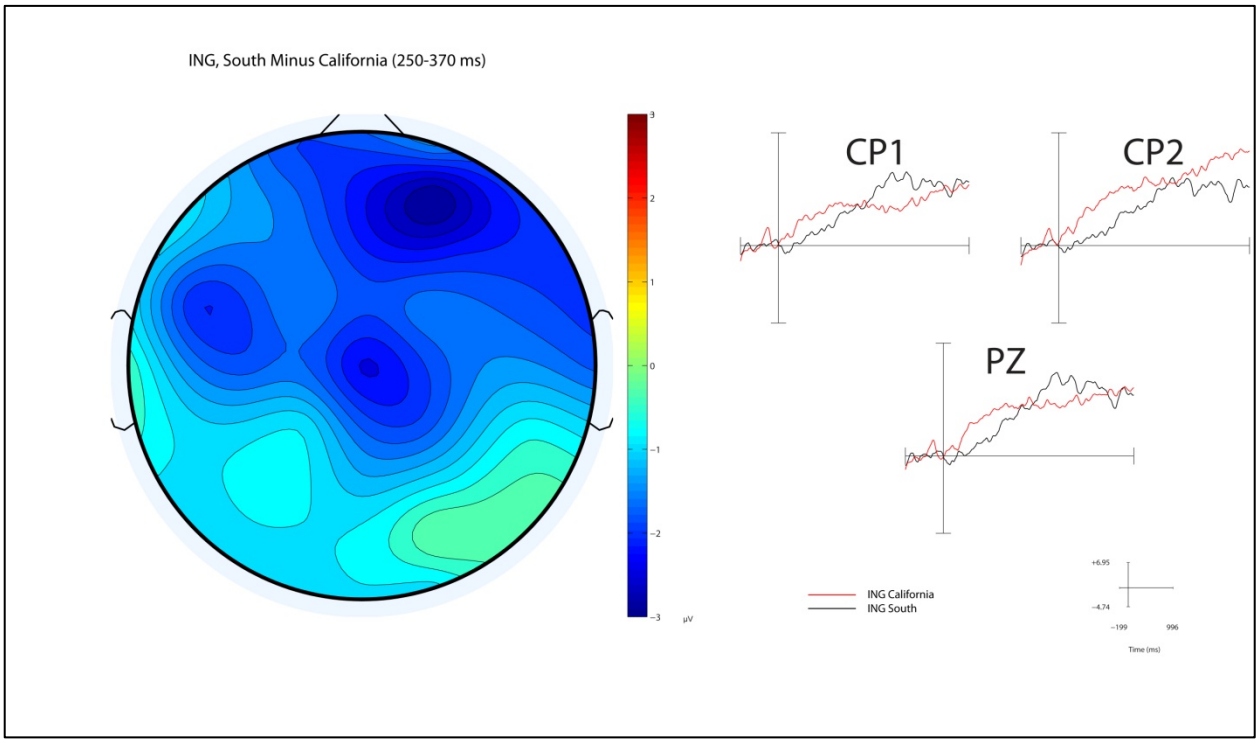

Figure 6.5 - Grand average ERP (n=19) difference ING, Southern minus Californian speakers (250-370 ms). Select electrode sites for ING critical words spoken by Caliornia speakers (red) and Southern speakers (black).

Nested variable analysis revealed the following significant pair-wise comparisons driving the Variant x Dialect interaction. For California speakers, Informal→IN' passages elicited waveforms -2.88 mV more negative than Formal→ING passages ($p = 0.020$). For the same California speakers, the Formal→IN' condition was 2.82 mV more negative than the Formal→ING condition. For the most canonical passages (i.e. Formal→ING), Southern speakers elicited waveforms 2.21 mV more negative than California speakers ($p = 0.023$).

6.5 Discussion

In this study we recorded Californian participants' event-related potentials while they listened to short passages containing ING/IN' variants spoken by Californian and Southern dialect speakers. With ERPs time-locked to the acoustic onset of the ING/IN' allomorphs we observed several modulations of the EEG signal within an N400 time-window: (1) a main effect of sociolinguistic variant between 260-510 ms, with IN' variants eliciting greater negativities than ING variants; (2) between 250 to 370 ms, dialect and variant interacted, with Californian, but not Southern passages eliciting an N400-like negativity for IN' compared to ING words – both the Informal→IN' and Formal→IN' conditions showing significantly greater negativities than the Formal→ING condition; and (3) Formal→ING passages elicited heightened negativities when spoken by Southerners compared to Californians.

In order to understand spoken language in real-time, psycholinguists concur that the brain must transform an acoustic representation of speech to a form that allows lexical access and retrieved word meanings must be integrated into higher-level structures. Questions remain, however, as to the nature and time-course of additional cognitive processes involved in language comprehension. In contrast to those who argue that there is an additional step after integration processes that makes use sociolinguistic information, our previous ERP experiment (Chapter 5) reported N400-like effects for sociolinguistic variants that were similar in latency, duration, and amplitude to those elicited by word cloze probability. Although we cannot discount the possibility that some effects of sociolinguistic information occur after lexical semantic processing, results from this first experiment are consistent with the proposal that both word meaning and social information can be integrated early and rapidly during real-time utterance comprehension. In ERP experiment #2, we sought to expand these effects by investigating the

role of speaker dialect. Here we sought to identify if the brain makes use information conveyed by speaker dialect to help predict which variable word form a speaker is likely to utter.

ERP experiment #2 replicated the primary effects observed in our previous work (Chapter 5). Between 260-510 ms, subjects showed an N400-like response to vernacular IN' variants that was similar in duration, amplitude, and distribution to that previously reported, but with a slightly delayed onset (i.e. 170 ms versus 260 ms). As with our first experiment, the Formal→ING condition elicited the greatest positivities compared to the Informal→IN' and Formal→IN' passages. However, in contrast to our previous study, we did not observe significant differences between the Informal→IN' and Formal→IN' passages. It will be recalled, however, that the significant differences seen in experiment #1 were only observed for critical words with low off-line predictability (i.e. cloze probability). The lack of low cloze words in the present experiment may factor in the inability to observe significant differences between these two conditions, as subtle psycholinguistic effects may only be observed when the processing system is taxed, for example frequency-based naming effects observed in immediate versus delayed naming (McRae, Jared, & Seidenberg, 1990). The consistency and timing of our observed effects, however, provide additional support that the observed electrophysiological response is due to processing the sociolinguistic aspects of speech rather than low-level perceptual or acoustic differences among stimuli.

The latency of this variant effect, within the temporal confines of the N400 window, suggests sociolinguistic variants can tax language processing in manner that is similar to that observed for words that are unexpected, semantically anomalous, or infrequent. Although there is controversy over what exactly the N400 indexes (For discussion see Lau, Phillips, & Poeppel, 2008), one interpretation argues that N400 waveform amplitude measures how well a word

integrates within larger conceptual representations of meaning (C. Brown & Hagoort, 1993; Hagoort & Van Berkum, 2007; Van Berkum, Hagoort, & Brown, 1999). In this conception, although sociolinguistic variants "mean the same thing," they differ in socially indexical meaning, and when this information conflicts with sociolinguistic expectations integration difficulties may arise (See Chapter 5 for discussion).

In the current experiment, we also observed an interaction of sociolinguistic variant and speaker dialect, namely increased negativities for vernacular IN' words compared to ING words when spoken by Californian speakers, but not when spoken by Southern speakers. Again, this finding is consistent with the effects reported in the previous experiment, which reported increased N400-like negativities for IN' variants compared to ING variants when read by Californian speakers (the only dialect speakers represented in experiment #1). In contrast, in the present experiment, Californian listeners hearing Southern speakers read the most canonical passages (i.e. Formal→ING) – passages that are incongruent with dialectal expectations for Southern speakers – elicited significantly increased negativities compared to Californian speakers. It should be noted that these observed interactions of dialect and variant which emerged around a typical N400 onset (250 ms), were relatively short-lived (~120 ms) in comparison to typical N400 effects (250-300 ms). Significant amplitude differences between conditions (2 - 3 mV), however, are within the range of typical N400 effects and are comparable to those we reported in our first experiment. Similar to the results observed in experiment #1, the N400-like potentials observed in this experiment showed a more frontal distribution in comparison to N400 effects of cloze probability and semantic anomaly.

Assuming these sociolinguistic effects are indeed related at least partially to N400 effects, these data suggest that language processing can be taxed by mismatches of speaker

dialect and sociolinguistic variant. When listeners hear speakers use a sociolinguistic variant that is atypical of that speaker's dialect, for example a Californian saying *workin'* or a Southerner saying *working*, language processing becomes more difficult. One possible explanation of these effects is that dialect speakers have difficulties pronouncing (ING) variants that are less common or less preferred in their own dialect. If these speakers have difficulties pronouncing dialect atypical variants, it could be possible that listeners are tuning to subtle disfluencies in pronunciation (e.g. stress, pitch, and prosody). One means of discounting this hypothesis would be to take acoustic measures of F0 and intensity for ING/IN' critical words in order to test if there are significant differences between speaker groups. Similarly, a behavioral experiment could test for speaker "naturalness" or "fluency" of ING/IN' words spoken by Californian and Southern speakers compared to non- ING/IN' control words.

An alternative possibility of these effects is word form plausibility. Perhaps, listeners find dialect atypical word variants implausible for specific dialect speakers. For example van Berkum and colleagues (2008), reported that sentences, such as "I cannot sleep without my *teddy bear* in my arms" were considered less plausible when spoken by an adult than when spoken by a child. In a previous behavioral experiment, which used the same stimuli read by six different California speakers, however, we reported no differences in plausibility rating for sentences with ING versus IN' words. Although ING sentences (mean 1.8) and IN' sentences (mean 2.1) were rated slightly less plausible than control sentences, no differences in ING and IN' sentence plausibility were observed [$t(56) = -1.71$, *n.s.*]. Although listeners found these ING/IN' sentences equally plausible for Californian speakers, it remains to be determined if subjects will rate these sentences differently for the speakers in experiment #2.

One possible explanation of these effects is that they are based on listener dialectal familiarity rather than dialect knowledge per se. Recall, that all of the subjects in ERP experiment #2 were native Californian listeners, and as such likely had greater exposure to and familiarity with Californian speakers compared to Southern speakers. If listening to unfamiliar dialect speakers is more difficult or more taxing than familiar dialect speakers, one might predict increased N400 negativities for the Southern speakers compared to the Californian speakers across conditions. Moreover, this account would predict that listeners hearing unfamiliar speakers using unfamiliar variants (i.e. Southerners speaking IN' variants) would elicit the largest N400 (perhaps analogous to the non-word N400 reported in other studies). In the current experiment, however, the only condition that elicited greater negativities for Southern speakers compared to California speakers was the Formal→ING condition. One future means of teasing out the effects of dialect familiarity would be to recruit two groups of subjects from different dialect backgrounds (e.g., Californian and Southerners) and exposing them to passages read by familiar and unfamiliar dialect speakers.

At this time we cannot entirely rule out effects of low-level differences among stimuli or effects of word form plausibility or dialectal familiarity. However, we believe the most parsimonious explanation of these observed effects, is that they reflect processing difficulties in integrating the sociolinguistic aspects of speech based on implicit dialectal knowledge cued by speaker voice. The onset of significant electrophysiological effects at 250 ms, outside the putative window of low-level acoustic processing, suggests that these effects are not entirely due to acoustic differences that may exist in these stimuli. Rather, the latency, amplitude, and polarity of the electrophysiological response, which is consistent with that reported in Chapter 5, suggests a taxing of higher-level cognitive processes involved in language comprehension. *These*

new results suggest there is nothing intrinsic about vernacular IN' variants that taxes language processing, but rather it is the combination of which variant is used and the dialect of the speaker that utters that variant and the context in which it is uttered that determines the processing cost.

In summary, it appears that the brain uses dialectal knowledge in order to predict what type of (ING) variant is likely to be uttered, similar to the way in which the brain makes use of lexical and syntactic information to help predict which word is likely to be uttered (e.g., cloze probability). One plausible interpretation of these results is that dialectal characteristics and their typical sociolinguistic variants are stored in long-term memory as a set of remembered exemplars based on previous sociolinguistic experiences. During real-time language processing, then, the incoming acoustic signal can be mapped to these stored exemplars during lexical access, pulling into working memory not just word meaning, but also sociolinguistic and dialectal information such as which variant is typical for a speaker of that dialect. One advantage of such a system, is that it would help facilitate on-line language processing by helping predict expected word forms, thus easing integration processes (For similar proposals, see Goldinger, 1998; Johnson, 2006). .

6.6 Summary

During human communication, language comprehension entails accessing word meaning and integrating it with information conveyed by the broader context. Our results show that information cued by sociolinguistic variation can be processed rapidly and early in the sense-

making process, suggesting few putative differences between word meaning and sociolinguistic information. Importantly, during this process the brain appears to make use of dialectal information conveyed by speaker voice to help predict which variable form is likely to be uttered. When listeners hear word variants that are incongruent with these regional dialectal expectations, integration processes may be taxed. This study provides evidence that during language comprehension, the language processing system is sensitive not just to *what* is said, but also to *how* it is said and *who* says it.

7. Listener Gender and (ING)

7.1 Introduction

In sociolinguistics an important area of investigation is the relationship between gender and language. As Bucholtz (2002) notes, within the variationist paradigm, the primary focus of gender and language has been in the study of sex differences between men and women; primarily with attention paid to quantifiable differences in production of particular, primarily phonological, sociolinguistic variables. Schilling-Estes (1998), for example, has argued that females use more standard language variants more frequently than males. One often noted observation is the apparent gender paradox - that women exhibit linguistic behavior that is both conservative and innovative (Labov, 1990, 2001a). Although women adopt innovative forms at higher frequencies than men, in terms of stable sociolinguistic variables, women demonstrate conservative behavior in their preference towards canonical and prescriptive forms. Labov (2001b) has codified these observations as a general principle of sociolinguistics: "For stable sociolinguistic variables, women show a lower rate of stigmatized variants and a higher rate of prestige variants than men."

Although considerable work has been conducted in this field, prior research has primarily focused on gender differences in *production* of variation and attitudes towards sociolinguistic variants. In contrast, relatively little is known about gender/sex differences in perceptual processing, though recent work is beginning to shed light on this matter. In the present chapter, I

examine gender differences in scalp-recorded brain activity while male and female subjects listened to passages that varied in (ING) realization and speaker dialect.

Over the past decades, a number of variationist studies have investigated (ING) realization by men and women in different speech communities. In early work conducted in a New England village, Fischer (1958) observed that females used the canonical velar ING variant at higher frequencies than their male counterparts. Less than a decade later, Shuy, Wolfram, and Riley (1967) in a study on Detroit speech made a similar observation, that females use ING more frequently than males. Trudgill's (1972a) study of speech in Norwich, England also made the same observation, as did Wald and Shopen's (1985) study in Canberra, Australia. Importantly, they also reported effects of *listener* gender on (ING) realization, observing that females increased IN' usage when speaking with males, suggesting that women "have a greater awareness of the stylistic effect of the style shifting" (p. 535). In general, these production studies show a general preference for canonical ING variants for women speakers compared to men. Table 7.1 presents a summary of gender and (ING) production studies.

Table 7.1 - Percentage of IN' by speaker gender. Adopted from Campbell-Kibler (2006).

Study	Location	Male	Female

Shuy et al. (1967)	Detroit	62 %	21 %
Shopen (1978)	Canberra	24 %	16 %
Houston (1985), <35 yrs	U.K.	88 %	72 %
Houston (1985) > 35 yrs	U.K.	78 %	76 %
Wald and Shopen (1985)	Canberra	23 %	24 %

Similarly, in terms of language attitudes, women often have a stated preference for the canonical ING form (Shopen, 1978). In recent language attitudes work, Labov et al. (2011) examined the effects of listener gender on attitudes towards (ING). Using a variation of the matched-guise technique, listeners were asked to rate the speech of a "newscaster-in-training" on a cline of occupational professionalism. The passages used as stimuli varied only in frequency of (ING) realization varying from 10 to 100% alveolar realization. Although both male and female judges showed similar logarithmic relationships between frequency and IN' realization, female listeners judged IN' use more negatively than their male counterparts (Fig. 7.1).

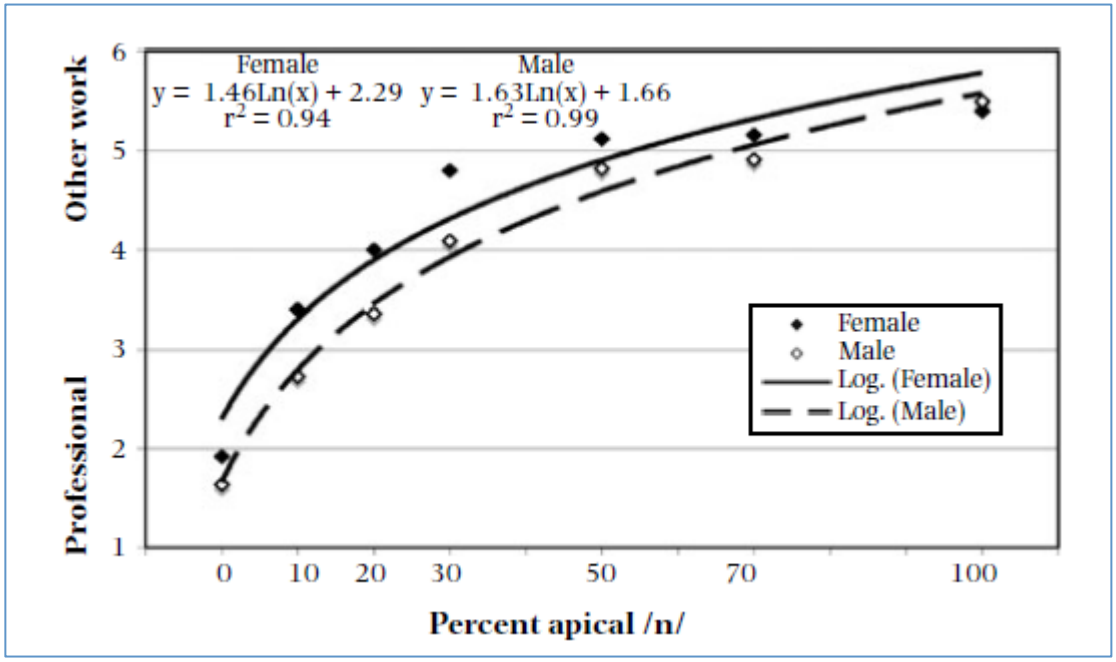

Figure 7.1 - Newscaster professionalism ratings by listener gender. Reproduced from Labov et al. (2011).

Most recently, the event-related potential technique has been used to investigate gender/sex differences in processing the social aspects of language use. In a study by van den Brink et al. (2010), male and female participants' ERPs were recorded in order to investigate gender differences in "the cognitive processes that mediate the integration of social information in a linguistic context." Participants in the study listened to sentences with messages that were either congruent or incongruent with stereotypical beliefs about the speaker, based on voice-based inferences about the speaker's socioeconomic status, age, or gender. For example, "I cannot sleep without my *teddy bear* in my arms" spoken by a young child versus an adult speaker; or "I have a large *tattoo* on my back" spoken in a posh accent versus a lower-class accent. Participants also listened to sentences that were either semantically felicitous or semantically anomalous (e.g., "You wash your hands with *soap/horse* and water"). As

illustrated in Fig 7.2, both males and females showed a large N400 to semantic violations (left), but only the female participants showed an N400 effect for the speaker identity manipulation (right). These results suggest that female listeners, but not males, are able to integrate information about the speaker with the linguistic message, making use of voice-based socially stereotypical inferences about the speaker during real-time language processing.

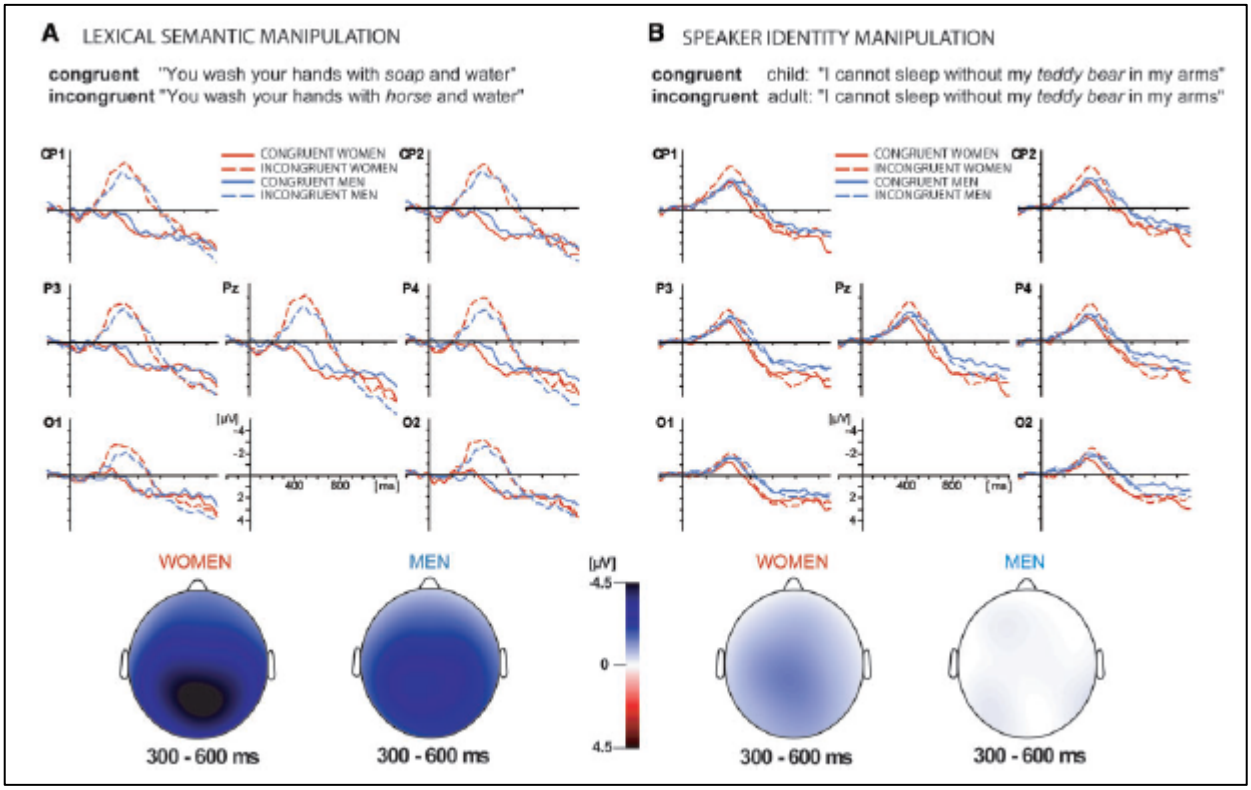

Figure 7.2 - Lexical semantic manipulation versus speaker identity manipulation. Reproduced from van den Brink (2010).

The present chapter adds to this body of literature by reporting gender-based electrophysiological differences in language processing. In the present study we reanalyzed the ERP data collected from two previous ERP experiments (Chapters 5 & 6) in order to investigate the role of listener gender in modulating the ERP response to sociolinguistic variation and

speaker dialect. In the two analyses presented we have added listener gender as a between groups factor in hopes of capturing some of the ERP variance reported previously and questioned whether female listeners compared to males might show greater sensitivity to the dialectal and sociolinguistic aspects of speech.

ERP Experiment #1, Listener Gender Analysis

In ERP experiment #1 (Chapter 5), Californian subjects listened to Californian speakers read passages manipulated for Register (formal/informal), Variant (ING / IN') and cloze probability (high/low). This experiment reported (1) an N400 cloze probability effect with low cloze words eliciting greater negatives than high cloze words; (2) an N400-like effect of sociolinguistic variant, with vernacular IN' variants eliciting greater negativities than canonical ING variants; and (3) an interaction of Variant x Congruency x Cloze x AP between 160-350 ms showing a hierarchy of N400-like negativities by condition, with Formal→ING and Informal→ING being the most positive, followed by Formal→IN', which in turn was followed by Informal→IN', the most negative condition. The observed hierarchy of ERP potentials by conditions, it will be recalled, best fit the predicted frequency account of sociolinguistic processing, though no significant differences were observed between ING critical words embedded in formal and informal register contexts as we had originally predicted (Table 7.2).

Table 7.2 - Observed versus predicted hierarchy in ERP experiment #1.

μV	Congruency Account	Frequency Account	Observed Hierarchy
+	Formal→ING	Formal→ING	Formal→ING = Informal→ING

	Informal→IN'	Formal→IN'	Formal→IN'
	Informal→ING	Informal→ING	
-	Formal→IN'	Informal→IN'	Informal→IN'

7.2.1 Gender predictions for experiment #1

Given the literature on gender/sex differences in variant production, language attitudes, and language processing, for this reanalysis we predicted an interaction of condition by listener gender. Specifically, we hypothesized that women compared to men should show greater sensitivity to vernacular variants and informal register contexts. This account predicts heightened N400-like negativities for females compared to males for critical IN' variants (i.e. Female(IN') < Male (IN')) and informal registers (i.e. Female(formal) < Male(informal)), but no gender differences for most canonical condition (i.e Female(Formal→ING) = Male(Formal→ING)). Following the results of van den Brink and colleagues (2010) we additionally predicted no significant gender differences by cloze probability (i.e. Female(low-high) = Male(low-high)). Importantly, the observation of a gender effect for sociolinguistic information in the absence of differences in cloze probability by listener gender, would provide additional evidence that the effects reported in Chapter 5 reflect the sociolinguistic aspects of language processing.

Gender Analysis for experiment #1

The previously acquired ERP data of 28 subjects (Chapter 5) was reanalyzed adding listener gender as a between groups variable. Subjects were divided into two equal groups of 14 men (18-23 years, mean age = 20.4 years) and 14 women (18-23 years, mean age 19.3 years). The resulting design, Variant (ING/IN') x Congruency (Congruent/Incongruent) x Cloze (High/Low),

by Gender (male/female) was analyzed in SPSS via repeated measures ANOVA with no correction for multiple comparison. We conducted two analyses with subjects' ERPs time-locked to word onset and the acoustic onset of the ING/IN' morpheme.

Results Experiment #1, Gender Analysis

In order to assess the role of listener gender in modulating the ERP response observed in experiment #1 (Chapter 5), we conducted two additional post hoc analyses adding listener Gender as a between groups factor. In the first analysis, we examined subject ERPs time-locked to word onset. As illustrated in Fig. 7.3, within the 230-510ms Cloze probability N400 window previously identified (Chapter 5), we observed no significant interaction of Cloze by Gender [$F(1,26) = 0.001$, $MSE = 8.696$, $p = .990$]. Both male and female listeners exhibited an N400 cloze probability effect, with low cloze words significantly more negative than high cloze words. As seen in Fig. 7.3, ERP waveform amplitude, latency, duration, and distribution for male and female listeners were extremely similar, exhibiting no significant differences in the aforementioned characteristics.

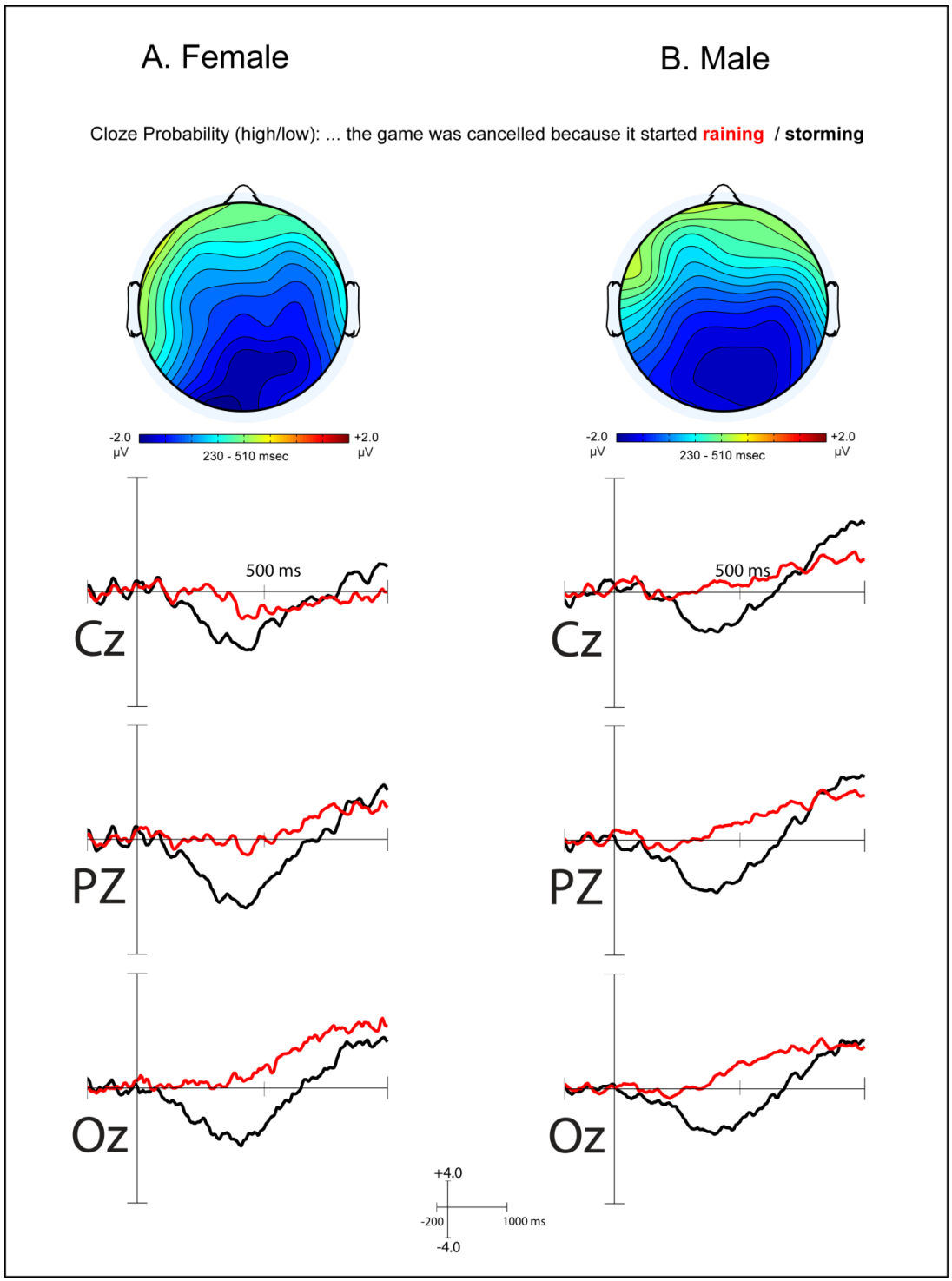

Figure 7.3 - Comparison of cloze effect by listener gender. Topo plots show grand average ERPs (Low Cloze minus High Cloze) 230-510 ms after word onset, for female (n=14, column A) and male (n=14, column B) listeners. Select midline plots show the effect of High Cloze (red) and Low Cloze (black) critical words (negative plotted down).

In the second gender analysis, we time-locked subjects' ERPs to ING/IN' morpheme onset, as previously discussed in the methods of Chapter 5. Here we observed an interaction of Variant x Congruency x Cloze x AP x LR by Gender between 160-350 ms [$F(1,26) = 5.15$, $MSE = 0.185$, $p = .032$], with females showing greater negativities than males for two conditions: Informal→IN' (1.63μV difference, $p=0.003$) and Formal→IN' (1.70μV difference, $p=0.046$) at posterior left electrode sites for high cloze words (Fig. 7.4). No additional significant pair-wise comparisons were observed.

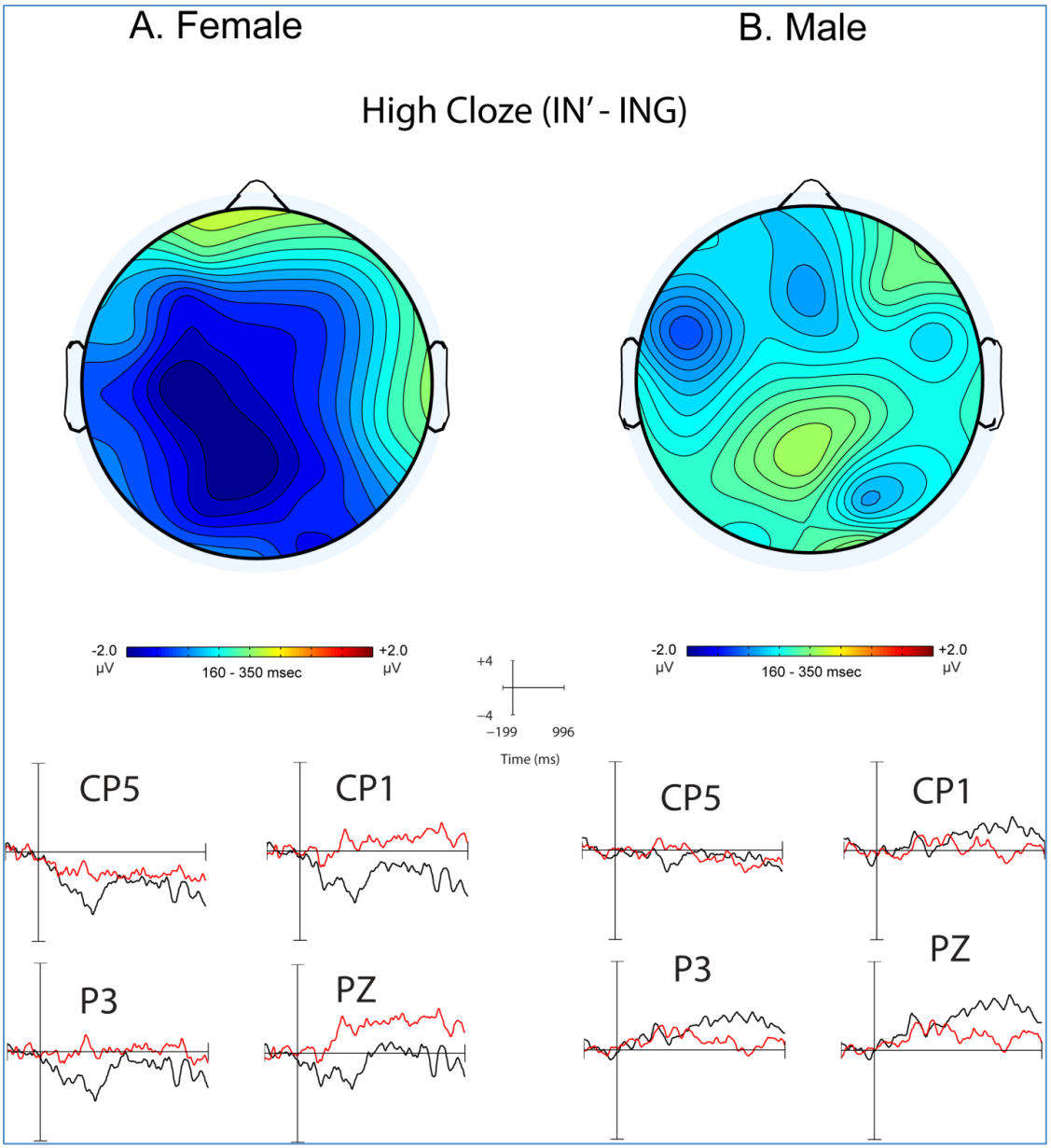

Figure 7.4 - Comparison of variant effect by listener gender for high cloze probability critical words. Topo plots show grand average ERPs (IN' minus ING) 160-350 ms after ING/IN' onset, for female (n=14, column A) and male (n=14, column B) listeners with ERPs locked to morpheme onset. Select plots show the effect of ING (red) and IN' (black) critical words at posterior left electrode sites (negative plotted down).

Discussion Experiment #1 Gender Effects

Although males and females showed no significant differences in N400 amplitude for cloze probability (Fig. 7.3), our analyses did reveal a Variant x Congruency x Cloze x AP x LR

interaction by listener gender, with females showing substantially greater N400 potentials for high cloze vernacular IN' variants compared to males (Fig. 7.4). Specifically, females exhibited N400-like negativities about 1.6 – 1.7μV greater than males for two conditions: Informal→IN' and Formal→IN'. We consider two explanations for these differences by listener gender: differences in auditory acuity and differences in sociolinguistic processing.

The first possibility is that the differential brain activity observed while males and females listened to these spoken passages is due to sex differences in auditory acuity and low-level auditory processing. Psychophysics studies on the perceived loudness of auditory stimuli, for example, have shown that women are more sensitive than men while listening to tones at different intensities (D'Alessandro & Norwich, 2009; Sagi, D'Alessandro, & Norwich, 2007) and that women have a lower comfortable loudness tolerance level than men (Elliott, 1971; McGuinness, 1974; Rogers, Harkrider, Burchfield, & Nabelek, 2003). Additional studies have shown auditory threshold differences by gender, with females showing superior auditory acuity (Dreisbach, Kramer, Cobos, & Cowart, 2007; Shahnaz, 2008). As there were non-significant trends in intensity for ING/IN' critical words, perhaps females' greater sensitivity to loudness contributed to these observed differences by listener gender. We offer two reasons why this proposal is unlikely. First, if gender differences in acuity were responsible, one would expect modulation of an earlier auditory component (N1), not an effect emerging 170 ms after the acoustic onset of the ING/IN' allomorph, outside the putative window of low level auditory processing. Second, if females' greater auditory acuity were driving these N400 variant negativities, one would have also expected similar gender differences in cloze probability, a result which was not confirmed by this experiment.

Although we cannot fully rule out that acoustic factors may be influencing aspects of these data, we believe a more viable explanation of these data is that they reflect gender differences in processing the sociolinguistic aspects of speech. As noted by Wolfram and Schilling-Estes (1998), for example, sociolinguistic studies of male-female language differences typically show that females tend to use more standard language variants than males. In consideration of variable (ING), in specific, Fischer (1958), Trudgill, (1974), Horvath (1985), and Labov (2001a), in studies conducted in the U.S., the U.K., and Australia, all report greater usage of vernacular IN' among men compared to women. Labov (2001b) has generalized these observations about gender and language variation as a general principle of sociolinguistics: "… women show a lower rate of stigmatized variants and a higher rate of prestige variants than men." In addition to using more prestige variants in formal registers, women often show an overall greater stylistic range than men (Schilling-Estes, 2002) as observed in a number of studies (e.g., Eckert, 1989; Labov, 1966, 1990). Moreover, as noted in the introduction, women have a stated preference for the ING variant and are less tolerant of vernacular IN' usage (Labov et al. 2011; Wald & Shopen, 1985). In summary these studies show, in terms of production and language attitude, women, compared to men, tend to use canonical variants at higher frequencies, show a greater stylistic range in variation, and hold more negative attitudes towards vernacular variants than men. We speculate that women, who use more standard variants in their speech and who show less tolerance for vernacular forms, are more sensitive to sociolinguistic norms.

Recent electrophysiological research speaks directly to this position. In an extension of earlier work, van den Brink et al. (2010) reported gender-based differences in brain activity while processing speech that violated socially stereotype-based inferences conveyed by speaker voice (e.g., an adult saying "I cannot sleep without my *teddy bear* in my arms"). Although

similar N400 components were observed for men and women while processing semantic anomalies (e.g., "You wash your hands in *horse* and water"), only the women showed N400 effects for passages that violated stereotyped-based inferences conveyed by the speaker's voice. The authors attribute these gender-based effects, not to sex-differences per se, but rather to individual differences in ability to empathize which typically correlates with gender. Although our study did not choose to investigate empathy, our gender data, which showed greater N400 potentials for variant but not cloze probability, are consistent with van den Brink et al, and suggest that women may use more automated and "deeper" sociolinguistic processing than men.

ERP Experiment #2, Gender Analysis

In ERP experiment #2 (Chapter 6), Californian subjects listened to Californian and Southern speakers read passages manipulated for Register (formal/informal) and Variant (ING / IN'). In Chapter 6 we reported (1) increased N400-like negativities for IN' stimuli spoken in a Californian dialect; and (2) N400-like negativities for canonical Formal→ING passages, spoken by Southern speakers compared to Californian speakers. Although these results generally fit with our predictions and with the results of ERP experiment #1, we did not observe the hierarchy of N400 potentials we had originally predicted as a function of register, variant, and speaker dialect.

7.2.2 Gender predictions for experiment #2

Keeping in line with our original predictions from Chapter 6, in this reanalysis of the data from ERP experiment #2, which included listener gender as a between groups variable, we predicted that females, but not males, would show greater sensitivity to passages that were *atypical* of speaker's dialect. Compared to males, we expected female listeners to show heightened N400-

like potentials for sociolinguistic variants that are incongruent with speaker dialect (i.e., IN' variants spoken by Californians and ING variants spoken by Southerners). We also predicted stronger variant register interactions for females, allowing us to observe the hierarchy of potentials predicted in Table 7.3.

Table 7.2 – Predictions for female listeners

	Californian Dialect	**Southern Dialect**	**N400 Amplitude**	**Interpretation**
Dialect Typical	Formal→ING	Informal→IN'	Smallest	Least Taxing
	Informal →ING	Formal→IN'		
	Formal→IN'	Informal →ING		
Dialect Atypical	Informal→IN'	Formal→ING	Largest	Most Taxing

7.3 Methods

7.3.1 Stimuli, Participants, and Design

The experimental design and methods are detailed in Chapter 6. We used the same EEG data previously collected from 19 native-speaker participants (female = 11) reported in Chapter 6. These participants ranged in age from 19 to 24 years and all were raised in Northern California (primarily the Bay Area and Central Valley).

7.3.2 Analysis

In order to investigate the role of listener gender in modulating the sociolinguistic N400-like electrophysiological response, we conducted a repeated measures ANOVA: Variant (ING/IN') x

Congruency (Congruent/Incongruent) x Dialect (Californian/Southern) by listener Gender (male/female), with no correction for multiple comparison. A full description of the methods, including data collection, filtering, and artifact rejection, can be found in the methods section of Chapter 6.

7.4 Results

Variant by Gender Effect, 100-180 ms (morpheme onset)

We observed a significant interaction of Variant by listener Gender between 100 – 180 ms after the acoustic onset of the ING/IN' morpheme [$F(1,17) = 7.034$, MSE 13.474, $p = 0.017$]. Planned pair-wise comparisons showed that IN' critical words increased negativities (1.06 mV mean difference) for female compared to male listeners ($p = 0.033$). For ING critical words, no significant differences in waveform amplitude were observed between gender groups ($p = .194$). Females listeners, but not males, showed increased negativities for IN' words compared to ING words ($p = 0.006$, 1.21 mV mean difference).

Congruency x Dialect by Gender Interaction, 200-350 ms (morpheme onset)

Between 200-350 ms after the acoustic onset of the ING/IN' morpheme, we observed a significant interaction of Congruency x Dialect x AP x LR by listener Gender [$F(1,17)=5.850$, MSE 0.532, $p = 0.027$]. Nested variable analysis identified the following pairwise comparisons were driving the Congruency x Dialect by Gender interaction. Informal→ING passages spoken by Californians

increased negativities for females compared to males at posterior right electrode sites (p = 0.036, 3.43 mV mean difference). Similarly, Informal→IN' passages spoken by Californians elicited greater negativities in females compared to males at anterior right electrode sites (p = 0.048, 3.39 mV mean difference). These gender contrasts are summarized in Table 7.4.

Table 7.4 - Experiment #2, Gender Analysis. Significant differences by listener gender.

Register→VARIANT	Speaker Dialect	Comparison	Quadrant	mV Difference	p-value
Informal→ING	Californian	Female < Male	PR	3.43	0.036
Informal→IN'	Californian	Female < Male	AR	3.39	0.048

For female listeners, but not males, significant differences by speaker dialect were also observed. Female listeners, hearing Informal→IN' passages showed greater negativities for Californian speakers compared to Southern speakers at posterior left (p = 0.008, 3.47 mV mean difference) and posterior right (p = 0.016, 2.87 mV mean difference) electrode sites (Fig. 7.5). Females listening to Formal→ING passages spoken by Southern speakers elicited more negative waveforms than Californian speakers at anterior left (p = 0.011, 3.39 mV mean difference) and anterior right (p = 0.012, 3.06 mV mean difference) electrode sites (Fig. 7.6). These significant differences by speaker dialect for female listeners are presented in Table 7.5.

Table 7.5 - Experiment #2, Gender Analysis. Significant differences for female listeners by speaker dialect. AL = anterior left, AR = anterior right, PL = posterior left, PR = posterior right. * p < 0.05, ** p < 0.01, t = trend, n.s. = not significant.

Register→VARIANT	Comparison	AL	AR	PL	PR
Informal→IN'	Californian < Southern	n.s	n.s	3.47 mV**	2.87 mV*
Formal→ING	Southern < Californian	3.39 mV*	3.06 mV*	n.s.	n.s

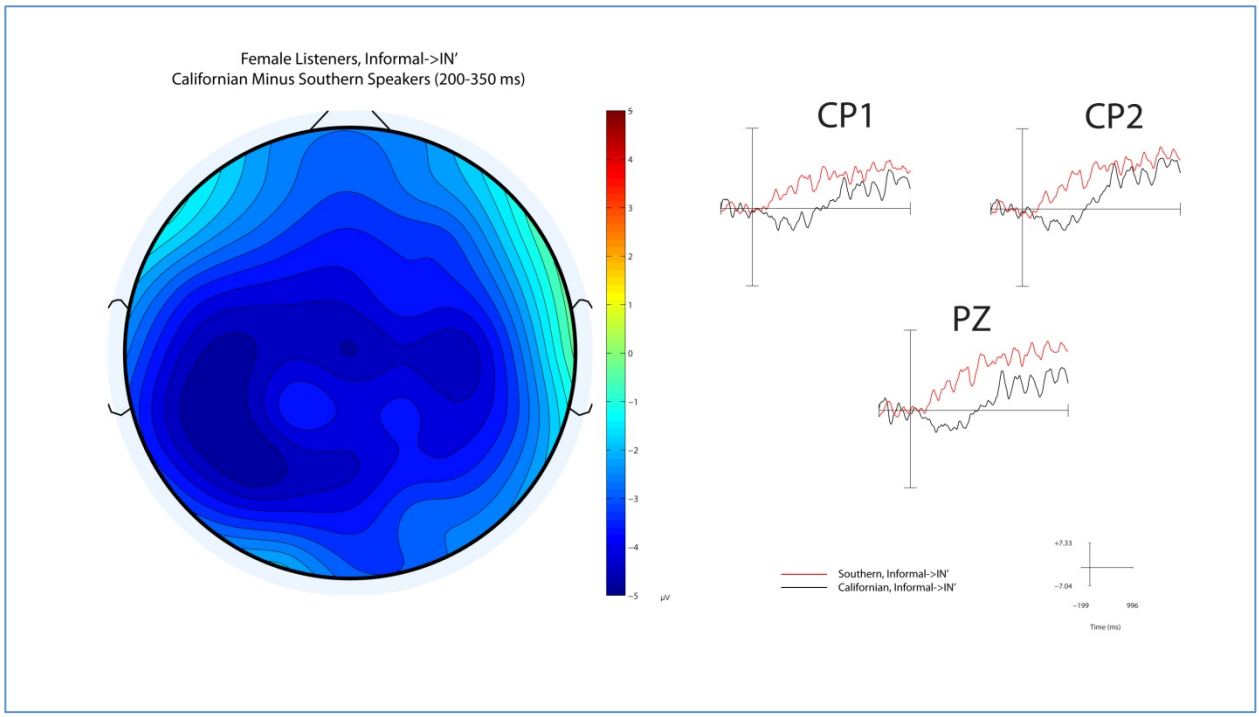

Figure 7.5 - Female grand average difference (n=11), Informal→IN', Californian minus Southern speakers (200-350 ms). Select electrode sites show Southern (red) and Californian (black) speakers.

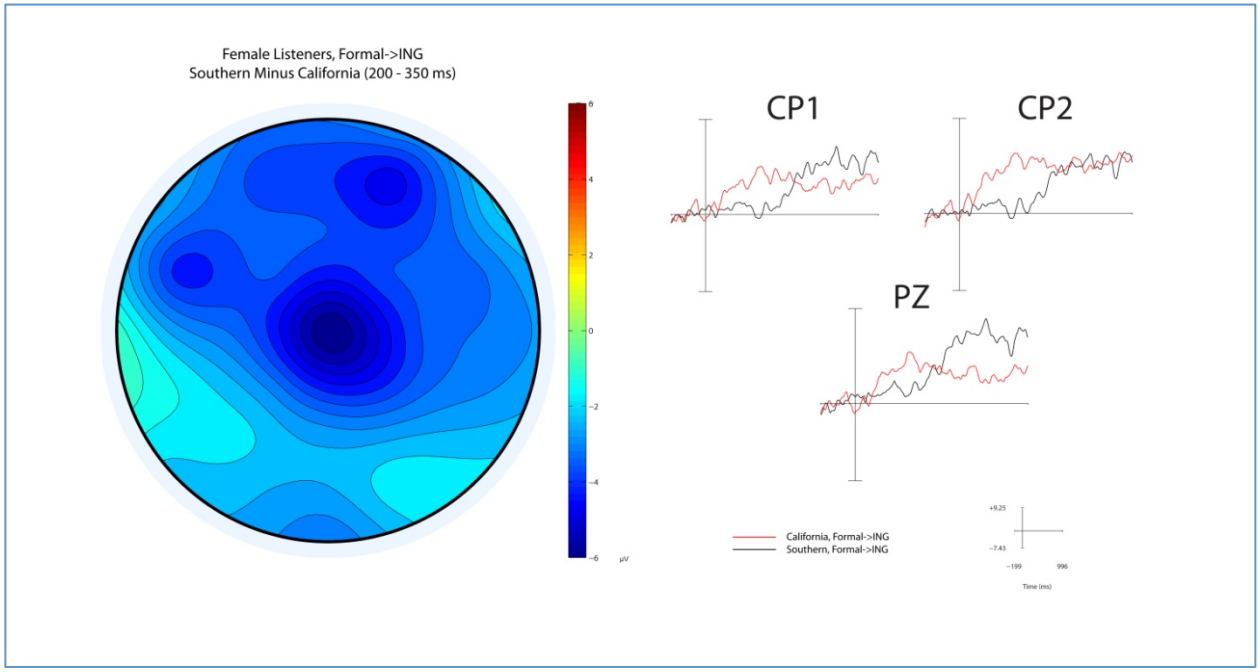

Figure 7.6 - Female grand average difference (n=11), Formal→ING, Southern minus Californian speakers (200-250 ms). Select electrode sites show Californian (red) and Southern (black) speakers.

In addition, female listeners, but not males, also showed significant amplitude differences between conditions holding speaker dialect constant. For females hearing Californian speakers, Informal→ING was more negative than Formal→ING at posterior right sites ($p = 0.025$, 3.65 mV mean difference, Fig. 7.7). For females listening to Southern speakers, Formal→IN' was more negative than Informal→IN' at anterior left electrode sites ($p = 0.040$, 2.14 mV mean difference, Fig. 7.8).

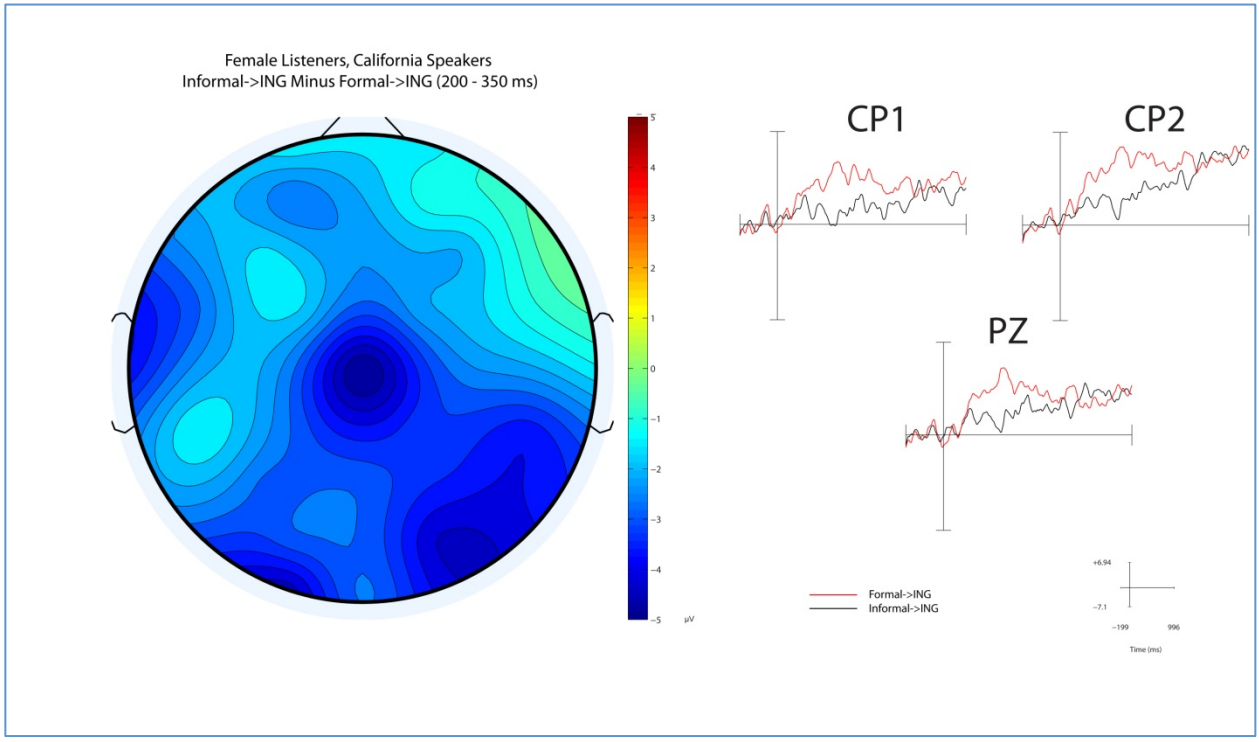

Figure 7.7 - Grand average difference, female listeners (n=11), California speakers: Informal→ING minus Formal→ING (200 - 250 ms). Select electrode sites show Formal→ING (red) and Informal→ING (black).

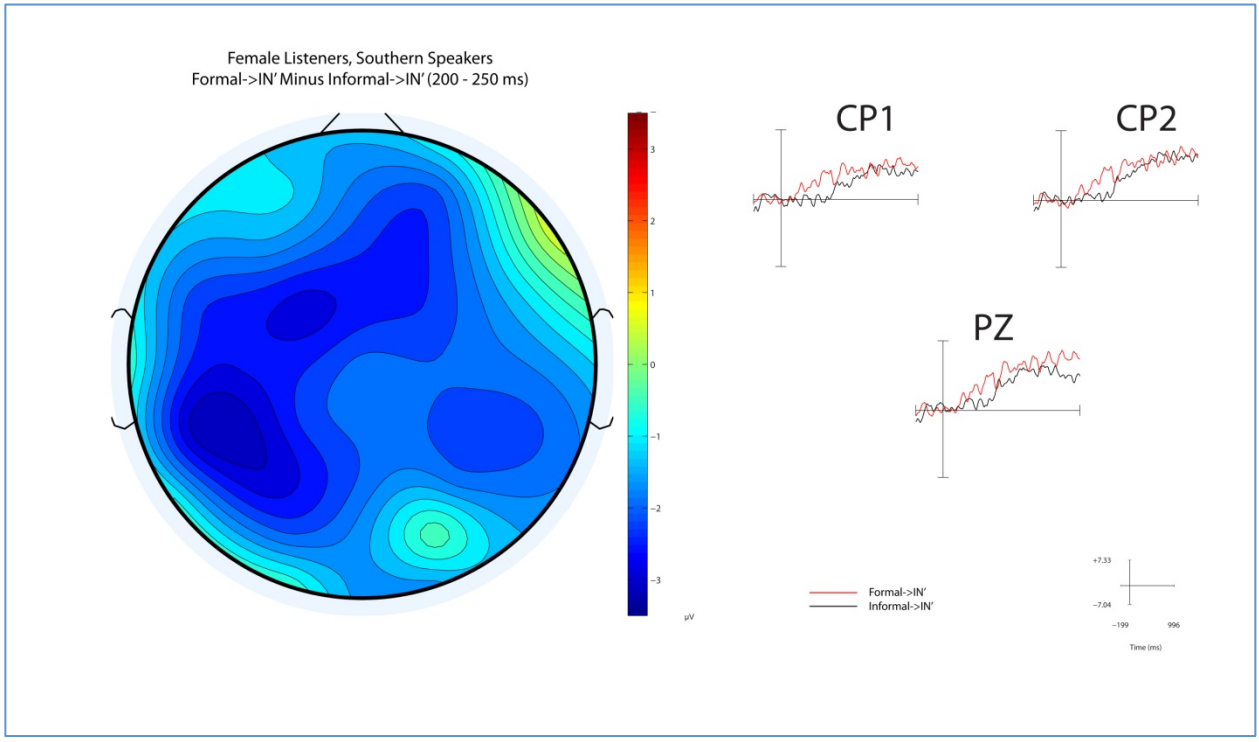

Figure 7.8 - Grand average difference, female listeners (n=11), Southern speakers: Formal→IN' Minus Informal→IN' (200 - 350 ms). Select electrode sites show Informal→IN' (red) and Formal→IN' (black).

For females listening to Californian speakers, Formal→IN' passages were significantly more negative than Formal→ING passages at anterior left ($p = 0.011$, 3.60 mV), anterior right ($p = 0.030$, 2.80 mV mean difference), posterior left ($p = 0.010$, 3.72 mV mean difference), and posterior right ($p = 0.022$, 3.35 mV mean difference) electrode sites. Similarly, we observed a statistical trend in females listening to Californian speakers for Informal→IN' passages to be more negative than Informal→ING passages ($p = 0.076$, 2.78 mV mean difference). Table 7.6 summarizes significant contrasts for female listeners.

Table 7.3 - Significant contrasts between conditions for female listeners by speaker dialect and electrode quadrant. AL = anterior left, AR = anterior right, PL = posterior left, PR = posterior right. * p < 0.05, t = trend, n.s. = not significant.

Speaker Dialect	Comparison	AL	AR	PL	PR
California	**Informal→ING < Formal→ING**	n.s.	n.s.	n.s.	3.65 mV*
Southern	**Formal→IN' < Informal→IN'**	2.14 mV*	n.s.	n.s.	n.s.
California	**Formal→IN' < Formal→ING**	3.60 mV*	2.80 mV*	3.72 mV*	3.35 mV*
California	**Informal→IN' < Informal→ING**	n.s.	2.78 mVt	n.s.	n.s

7.5 Discussion

In this exploratory reanalysis of the ERP data from experiment #2 (Chapter 6), we added listener gender as a between groups factor. We report several novel findings: (1) between 100-180 ms, variant and listener gender interacted with females but not males showing increased negativities for IN' compared to ING variants; (2) between 200-350 ms, females showed an N400-like negativity for Informal→IN' passages spoken by Californians and for Formal→ING passages spoken by Southern speakers; (3) females showed heightened negativities for Californian ING variants preceded by an informal compared to a formal context, and heightened negativities for Southern IN' variants when preceded by a formal, compared to an informal register context; and (4) females showed heightened negativities for **IN'** variants compared to **ING** variants when spoken by Californian speakers.

Based on the gender differences reported in other literature and those observed in ERP experiment #1, we had predicted that female listeners compared to males would show greater sensitivity to the sociolinguistic aspects of speech which would be reflected in heightened N400-like negativities to word variants and register contexts that were *atypical* of speaker dialect. Between 200-350 ms, females, but not males, showed increased negativities for the most dialect

atypical passages. Californian speakers, who preferentially use **ING** variants in their own speech (see Chapter 5), when uttering the *most dialect atypical* passage (i.e. Informal→IN') elicited significantly greater negativities compared to these same passages when read by Southerners. Similarly, Southern speakers, who preferentially use **IN'** variants in their own speech, when uttering their *most dialect atypical* passages (i.e. Informal→IN') elicited significantly greater negativities compared to these same passages when read by Californians. In contrast, male listeners showed no such effects. These findings suggest that female listeners primarily drove the significant aggregate-level effects previously reported in Chapter 6.

Importantly, however, it wasn't just the *most* dialectal atypical passages (i.e. Californian Informal→IN' and Southern Formal→ING) that increased negativities. Female listeners, but not males, also showed increased negativities for *register contexts* that were inconsistent with speaker dialect. Listening to Californian speakers utter a dialect typical ING variant preceded by a Californian *atypical* informal register context elicited heightened negativities compared to a Californian *typical* formal register context. Similarly, listening to Southern speakers use a dialect typical IN' variant preceded by a Southern *atypical* formal register context elicited heightened negativities compared to a Southern typical informal context. That is to say, hearing Californian speakers use an informal register context and hearing Southern speakers utter a formal register context both elicited heightened negativities compared to their respective dialect typical contexts. This pattern of results (Table 7.7) and the significant differences by speaker dialect for dialect atypical passages is consistent with our earlier predictions (Table 7.2).

Table 7.4 - Dialect atypical register effect.

	Californian Dialect	**Southern Dialect**	**N400 Amplitude**	**Interpretation**

Dialect Typical	Formal→ING	Informal→IN'	Smaller	Less Taxing
Dialect Atypical	Informal →ING	Formal→IN'	Larger	More Taxing

Although we cannot entirely rule out sex differences in acuity (see earlier discussion, this chapter), and future work will be needed to evaluate to the statistical significance of these effects given the small sample size (n=19), we believe these results are consistent with those reported in the gender analysis for ERP experiment #1, those reported by van den Brink et al. (2010), and those reported in the sociolinguistic literature (e.g., Labov et al., 2011). Specifically, we believe this pattern of data reflect gender differences in processing the sociolinguistic aspects of speech. Across the two gender analyses, we summarize these observations: (1) no gender differences in N400 amplitude for word cloze probability – both males and females showed cloze probability N400s of similar magnitude, latency, duration, and morphology; (2) female listeners showed greater N400-like negativities compared to males while listening to Californian speakers utter passages that were *entirely* dialect *atypical* (i.e. Informal→IN', exp.#1 & exp.#2) or *partially atypical* (Formal→IN',exp.#1 & Informal→ING, exp.#2) – in contrast, no gender differences were observed in either experiment for Californian speakers uttering the *most dialect typical* Formal→ING passages; and (3) in experiment #2, females, but not males, showed amplitude differences for speaker dialect – for Informal→IN' passages, Californian speakers were more negative than Southern speakers and for Formal→ING passages, Southern speakers were more negative than Californian speakers. These female listeners also showed increased negativities when speakers used preceding register contexts that were inconsistent with their dialect.

The time course of the ERP modulation (~ 160-350 ms) and the consistency of these results with previous gender ERP studies (e.g. van den Brink et al., 2010), production studies, and language attitude studies provide additional support for the claim that females appear

particularly sensitive to manipulations of sociolinguistic aspects of speech. When these sociolinguistic and dialectal expectations are violated during language processing, female listeners incur a processing cost as reflected by increased N400-like negativities. Importantly, for these female listeners, it wasn't just the most egregious violations of speaker dialect, variant, and register that elicited these processing costs. Females also showed particular sensitivity to the preceding register context. Californian speakers using formal contexts and Southern speakers using informal contexts eased processing in comparison to the speakers using dialect atypical contexts. In summary, female listeners compared to males, show heightened effects when encountering variable word forms and register contexts that "violate" probabilistic sociolinguistic and dialectal expectations.

As previously discussed, variable (ING) is a stable, stigmatized sociolinguistic variable, that is to say a variable which is not undergoing linguistic change and which overtly and accurately associated with informality. As Eckert (1989) notes, women "tend to be more conservative than men in their use of those vernacular forms that represent stable social variables" (p. 247). A number of theories have been proffered to explain this general pattern. Trudgill's study of Norwich English characterized women as generally conservative, as men led most of the changes in progress, arguing that women oriented towards overt standard-language prestige norms, whereas men oriented covert, vernacular prestige norms (Trudgill, 1972b). It has also been argued that women's conservatism in speech reflects gender power relations in society, with the use of standard language forms functioning as a face-saving strategy in un-egalitarian contexts (Deuchar, 1988). Labov (1984) has argued that women orient towards global prestige norms, refraining from using variants that are stigmatized in the larger community (i.e. stable sociolinguistic variables and older, stigmatized changes in progress).

Whatever the reasons that women show conservative linguistic behavior with stable sociolinguistic variables, one question that emerges is *how* do they actually do this? Although speculative at this point in time, one tantalizing prospect is that perhaps it is this heightened sensitivity to the social dimensions of speech that allows women to tune their own production patterns accordingly according to sociolinguistic norms. Such a proposal, for example, supports Wald and Shopen's (1985) observation that females increased IN' usage when speaking with males, bolstering their claim that women "have a greater awareness of the stylistic effect of the style shifting" (p.535). Future electrophysiological work will need to investigate this further by studying additional sociolinguistic variables, including variables undergoing linguistic change as well as variables that are not overtly stigmatized.

7.6 Summary

The present event-related potential study testifies to the importance of investigating gender differences in language processing, especially differences in processing the social aspects of human communication. By considering listener gender, what was considered noise in our aggregate analyses was removed, revealing important differences in brain activity between males and females while processing the sociolinguistic and dialectal dimensions of language use. Specifically, we found that female listeners showed particular sensitivity to conditions where sociolinguistic variants clashed with expectations engendered by speaker dialect. Under these conditions, females showed large N400-like negativities that were absent in the males. These findings join a small but growing body of literature that supports the contention that women may use more automated and "deeper" sociolinguistic processing than men, suggesting greater sensitivity to the social dimensions of language.

8. Conclusions: Perception of Sociolinguistic Variation

This dissertation sought to extend our knowledge of sociolinguistic cognition – the goal of which is to characterize the computational stages and cognitive representations underlying the perception and production of sociolinguistic variation. In particular, this project attempted an initial characterization of the processes that support the perception of spoken word variants as well as to identify some of the salient contextual factors that modulate these operations. Specifically, this project which focused upon the single sociolinguistic variable (ING), used the event-related potential technique to characterize the time course and processing costs associated with encountering sociolinguistic variants during real time language comprehension.

As one of the first electrophysiological studies examining the perception of sociolinguistic variation, much of the novel work in this dissertation was necessarily exploratory in nature. Although perception of variation and language attitudes towards variation has been central to variationist sociolinguistics, much of this early work has been somewhat limited by the methodologies and measurements that were used. Measurements such as the matched-guise technique, where participants rate speakers via semantic differentials (e.g., Campbell-Kibler,

2007; Labov, et al., 2011), are off-line measures which pose severe limitations for studying real-time, dynamic language processes. Other off-line measures, such as forced and free choice auditory classification of dialects (e.g., Clopper & Pisoni, 2004; Clopper & Pisoni, 2007) are limited by ecological validity. Although individuals can and do (at least sometimes) make judgments of speaker origin and speaker traits, one wonders how reflective these tasks are normal, real time, natural language comprehension.

The present study attempted to overcome some of these previous limitations by using real-time measurements of brain activity under (relatively) ecological conditions with naturalistic task constraints. In these experiments, participants listened to short, semantically rich stories – stories not much different from the type that are told around the dinner table – about little league baseball games, family vacations, and Sunday afternoons in the park. Similar to the speaker diversity encountered in daily life, these passages were spoken by different speakers, some male, some female, some from California, some from elsewhere. Moreover, rather than have participants make overt judgments of passages or speakers, these experiments merely required participants to listen for comprehension, one of the most natural tasks possible for studying language processing under ecological conditions.

Only by using such precise measures as event-related brain potentials, which have exquisite temporal resolution at the millisecond level, can we begin to understand and characterize the cognitive operations involved in real time comprehension of spoken language variation. As this was an exploratory investigation, (ING) was an excellent variable to study because it is frequent in natural conversation, all English speakers show variability in its pronunciation, it is a sociolinguistic stereotype, and it has been well-studied from the perspective of language production and language attitudes. This later point was particularly crucial for the

current experiments because it facilitated predictions of brain activity based on the rich literature on (ING) developed by sociolinguistics over the past five decades.

Summary of Findings

In the present section, I summarize the primary findings reported in this dissertation.

ERP Experiment #1 (Chapter 5) tested whether event-related potential measurements would be sensitive to spoken-language variation and sought to characterize the nature and time course of cognitive processes involved in real time language comprehension. In this experiment, participants ERPs were recorded while they listened to short passages read by different Californian speakers and answered true/false questions about the passages they heard. Passages were constructed to engender strong expectations for sentence-final critical words (i.e. cloze probability) and were additionally manipulated to vary by (ING) variant and preceding register context. Results showed (1) a classic N400 effect of word cloze probability (low cloze minus high cloze); (2) a similar N400-like effect of variant (IN' minus ING); and (3) a hierarchy of N400-like negativities across condition (Formal/Informal→ING > Formal→IN' > Informal→IN').

According to the results of experiment #1, it was argued that the brain begins to integrate word meaning and sociolinguistic indexical meaning around 170-210 ms after acoustic onset of the word or variant. The results showed that during normal spoken language comprehension, brain processes are sensitive to sociolinguistic variants that can serve as cues to social information. The temporal characteristics of the N400-like brain response variation further suggested that sociolinguistic variation is rapidly brought to bear in the on-going sentence interpretation. The similarity of brain response to manipulations of cloze probability and sociolinguistic variant, suggest there may be few differences between formal semantic meaning

and social meaning conveyed by how a word is pronounced. Slight differences in distribution, however, suggest different cortical structures are involved in these processes, and future research will need to further characterize these differences. The results from this first experiment were interpreted as support for a single-step model of language processing, one that can rapidly take into account sociolinguistic aspects of speech during language comprehension.

ERP Experiment #2 (Chapter 6) sought to replicate and extend the findings of the first experiment. Specifically, this experiment sought to broaden the scope under which sociolinguistic information elicits N400-like negativities. In our earlier experiment we had observed an N400 effect of variant – IN' variants elicited heightened negativities compared to ING variants. Here, in the second experiment, we sought to reverse these effects by introducing the variable of speaker dialect. In experiment #2, subjects listened Californian and Southern dialect speakers, read passages containing ING /IN' word variants in formal/informal register contexts. The question this manipulation addressed is whether listeners' use information conveyed by speaker dialect to predict which (ING) variant or register the speaker is likely to utter? Between 250 to 370 ms, we reported a dialect and variant interaction, with Californian, but not Southern passages eliciting an N400-like negativity for variant, with the Informal→IN' and Formal→IN' conditions showing significantly greater negativities than the Formal→ING condition. In this same time window, Formal→ING passages spoken by Southerners were significantly more negative than when these passages were read by Californian speakers. These results were taken as evidence that the listener's brain can make use of dialectal information conveyed by speaker voice to help predict which variable form a speaker is likely to utter. During language comprehension, when listeners encounter word variants that mismatch these dialectal expectations, processing can be taxed. Taken together, the results from these two ERP

experiments provide compelling evidence that during natural language comprehension, the language processing system is sensitive not just to *what* is said (cloze probability), but also to *how* it is said (ING / IN') and *who* says it (Californian/Southerner).

In the gender analyses of Chapter 7, we reanalyzed the ERP data from the previous two experiments. Adding listener gender as a between-groups variable, these analyses attempted to capture group-level gender differences in processing the sociolinguistic aspects of speech. Both analyses reported similar gender-based differences in scalp-recorded brain activity. Experiment #1, showed no male/female differences for word cloze probability, but female listeners did show heightened N400-like negativities for vernacular IN' variants compared to male listeners. In Experiment #2, females, but not males, showed N400-like negativities for passages that were incongruent with dialectal expectations, for example, Californians reading passages with IN' variants, or Southerners reading passages with ING variants).

These gender results are largely consistent with other studies examining gender and variation from the perspective of language production (e.g., Fischer, 1958; Shuy, Wolfram, & Riley, 1967), language attitudes (e.g., Labov, et al., 2011), and the event-related potential technique (e.g., van den Brink, et al., 2010). Although more work needs to be conducted in this domain, these initial findings suggest that women may use more automated and "deeper" sociolinguistic processing than men – a proposal that likely resonates with many varaitionists.

Sociolinguists have long-recognized variable word forms differ in social correlates and social meaning. The results from this study demonstrate that social meaning impacts language processing early in the comprehension process in a manner analogous to word semantics. Because the listener's brain treats variants differently during language processing, this suggests a tacit level of sociolinguistic knowledge that can be rapidly deployed in the service of real-time

spoken language comprehension. During language comprehension (Fig. 10.1), listeners appear to make use of dialectal information conveyed by speaker voice in order to help predict which variable form is likely to be encountered. When listeners hear sociolinguistic variants that "violate" these dialectal expectations, language processing is taxed. The results presented in this dissertation have hopefully provided some compelling evidence that during language comprehension, the brain is sensitive not just to *what* is said, but also to *how* it is said, *who* says it, and *who* hears it.

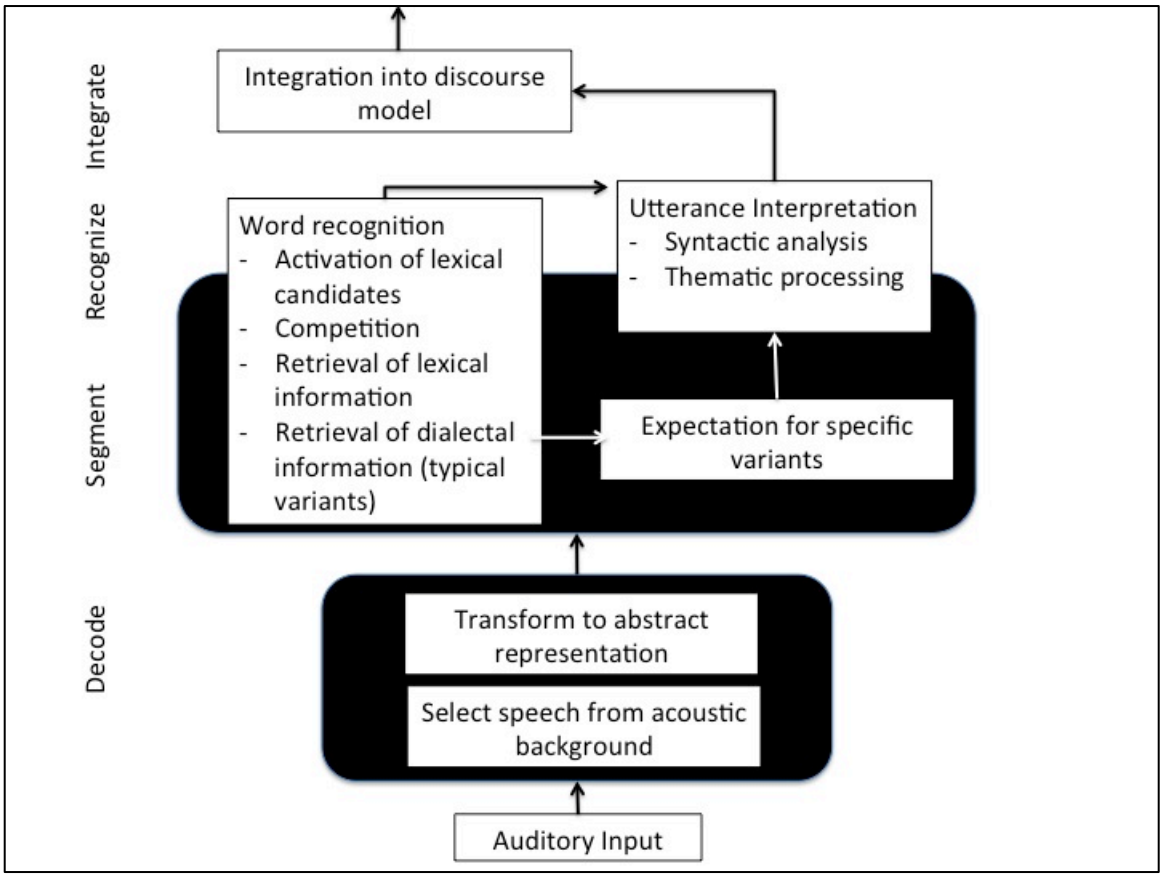

Figure 8.1 - Sociolinguistic model of language processing

Limitations and Future Directions

One of the primary limitations of this project was the subject sample – all of the listeners were WEIRD, selected from Western, Educated, Industrialized, Rich, and Democratic (WEIRD) societies, specifically American, undergraduate students enrolled in psychology coursework (Henrich, Heine, & Norenzayan, 2010). As such, questions remain as to how representative these subjects' behavior and brain activity are to a more general population of English listeners. Decades of sociolinguistic research, for example, have identified speaker age, education, and SES as contributing factors to variation production. Whether a change in progress or age graded stable variation, addressing participant age (as well as additional factors such as socio-economic status) has been central to uncovering the systematic regularities that govern sociolinguistic variation. Future work in sociolinguistic cognition will need to move out of the university laboratory and out into the "real world" with "real" subjects, who are likely to show differential patterns of production, perception, and attitudes.

The present study only recruited university students enrolled in a large, California university. This poses a limitation not just on the generalizability of these findings, but also on their interpretation as well. For example, experiment #2 (Chapter 6) reported a dialect variant interaction, with increased N400 negativities for dialect speakers using atypical variants (i.e., Californian IN' and Southern ING). Because only California listeners were used, it is unclear whether the increased negativities for Southern speakers using ING variants is due to the "Southerness" of the speaker or the "not my dialect" of the listener. One way future work will start addressing these issues is by recruiting subjects from different regions, as well as participants differing by age, gender, and socio-economic status.

Another limitation of this study is the type of stimuli participants listened to. Specifically, subjects listened to artificially constructed passages read by different speakers. As others have

observed, conversational speech compared to read speech is characterized by more frequent use of vernacular variants (e.g., Labov, 1966) and naive listeners have few difficulties distinguishing between spontaneous spoken and read speech (Blaauw, 1994; Laan, 1997). As such, some of the effects observed may be due to this juxtaposition of a relatively formal register (i.e. read scripts) using informal vernacular variants. A more ecologically valid alternative for future research is to examine spontaneous, extempore speech, although this poses obvious methodological hurdles such as controlling likely influential linguistic variables (e.g., word frequency, critical word sentence position, etc.). One possible solution that balances experimental control and ecological validity was presented by Campbell-Kibler (2006). Specifically, she collected extempore speech elicited through sociolinguistic interviews, and then had the original speakers return to the studio to record the same passages in different (ING) guises. Only by examining naturalistic speech – spoken by real speakers in diverse communities – will we begin to truly understand the neural bases of sociolinguistic cognition.

In the following paragraphs, propose a number of likely fruitful avenues for future research investigating sociolinguistic cognition in general and variable (ING) in particular. These experiments, which include staple psycholinguistic techniques such as chronometric priming experiments, as well as ERPs, are designed to address some of the limitations of the current study, including more carefully characterizing processing costs, variation representation, dialectal exposure, and the neural correlates of sociolinguistic cognition.

One of the first things that need to be further characterized is the short and long-term processing of, and representation of (ING). Following work by Sumner and Samuel (2005, 2009), this proposed experiment is designed to investigate how dialectal variants are processed and represented over different time spans. In order to determine the immediate and long-term

processing consequences for encountering (ING) variants, an auditory semantic priming paradigm will be employed. In this paradigm participants listen to prime-target word pairs and make a lexical decision on the target word. The prime word *running* for example should prime, or speed the reaction time, for the related target word *jogging* compared to the unrelated target word *throwing*. In this experiment, participants will hear prime words consisting of two regular (ING) variants of a word (e.g., canonical *running* and vernacular *runnin'*). Target words are semantically related or unrelated to the prime, and either match or mismatch the (ING) variant of the prime. For example, given the prime *running*, the possible targets are: semantically related matched variant – *jogging*; semantically related mismatched variant – *joggin'*; semantically unrelated matched variant- *throwing*; semantically unrelated mismatched variant – *throwin'*). Results from this experiment will determine if there are immediate processing costs associated with processing non-canonical regular (ING) variants.

The second proposed experiment uses a long-term repetition priming paradigm in which all prime words are presented in a first block of trials and all target words are presented in a second block of trials. This experiment differs from the first, in that it also uses a repetition priming paradigm in which target words are repeats of prime words. Using this design will allow researchers to investigate the degree to which one (ING) variant primes another variant relative to itself. The temporal lag in the blocked design should enable identification of which form is encoded in memory over the intervening interval. Results from this experiment should determine whether the canonical or vernacular form serves as the underlying representation.

A final priming experiment will investigate the effects of dialectal experience on the perception and representation of (ING) variation. In this between groups study, participants will be recruited from California and the Southern United states. Whereas most California speakers

predominantly use the ING variant (Chapter 5 & Chapter 8), most Southern speakers use the IN' variant in their speech (Labov, 2001). This study proposes to use both the short-term and long-term discussed in the previous paragraphs, in order to better characterize how exposure to (ING) variants relates to underlying form of these different populations of dialect speakers.

An additional experiment proposes an extension of our ERP paradigms presented in Chapters 5 and 6. The reader will recall from these earlier studies, that the predicted hierarchy of N400 potentials was only partially observed. One speculation is that these predictions were only partially manifest because of the random design used in those experiments (i.e. all trials were completely randomized for each subject for each of the eight lists). This experiment proposes a fully blocked design, where each block will consist of a single condition (e.g., formal register context with high cloze canonical ING variant). The block design is predicted to be sensitive enough to measure differences between conditions that were not apparent in the previous two ERP experiments.

For future research in sociolinguistic cognition several questions that still remain can be addressed through variations of the (ING) ERP paradigms:

1.) Do listeners take into account the gender of the speaker while processing sociolinguistic variation? The production literature consistently shows female speakers are more conservative in their use of stigmatized, stable sociolinguistic variables. This predicts heightened N400 negativities listening to vernacular IN' variant spoken by female speakers compared to males.
2.) Are stable and change in progress sociolinguistic variants processed equally? Variable (ING) is a stable variable, but others reflect a change in progress, for example, the

Northern Cities Chain Shift. Would females, the typical leaders of linguistic change in progress, show a reversal of the gender effects we observed in Chapters 5 and 7?

3.) Are there differences in processing costs for encountering stereotypical variables compared to sociolinguistic markers and indicators (Labov, 1971)? English (-t,d) deletion, for example, though just as widespread and stratified as (ING), doesn't appear to be as stereotypical – naïve speaker/listeners may be unaware of its presence and may not have direct conscious access to any implicit attitudes associated with this variable. Does encountering a (-t/d) deleted form in spoken discourse elicit the same N400-like negativity as vernacular IN'?

4.) Does the lexical class of the word and the local phonological context determine the processing cost of sociolinguistic variation? Numerous variationist studies have identified the internal linguistic factors that constrain the production of sociolinguistic variants (e.g., the verbal/nominal conditioning of (ING), the phonological conditioning of (-t/d) deletion, etc.). Does the listener take into account these factors to help predict which variant is likely to be realized in the speech of others?

The above questions are just a few of the many that still remain in the nascent field of sociolinguistic cognition, which will need to be addressed in the coming years.

Conclusion

This chapter has presented some of the theoretical implications and methodological limitations of this dissertation project. It has argued that listeners have mental representations of sociolinguistic variation, and that these representations are used in the service of natural, real-

time language comprehension. I have also provided a number of fruitful avenues for future research using both chronometric and electrophysiological measures to investigate the processes and neural correlates of sociolinguistic cognition.

This study, focused upon the single sociolinguistic variable, English (ING), and investigated the how the perception of these spoken variants modulates brain activity during language comprehension. It has argued that listeners have fine-grained representations of sociolinguistic variation which includes information about regional dialect sociolinguistic preferences. The results from these ERP experiments have shown how the brain/mind uses these representations in the service of language processing by helping predict which variable form a speaker is likely to utter. Importantly, this study also identified how these processes are further modulated by individual differences in listener gender, implicit stereotypical attitude towards variation, and listener conservatism in their own speech. In summary, this work provides strong evidence that the cognitive mechanisms that support language comprehension depend on *what* is said, *how* it is said, *who* says it, and *who* hears it.

A

Examples of stimuli. The context contained 4 words inflected with progressive morphology (bold) which were read in two different guises, formal (all ING pronunciations) and informal (all IN' pronunciations). The sentence-final critical words were high or low cloze and were spoken with either the ING or IN' pronunciation. Critical words were read and spliced in a phrasal context (underlined).

Context	High Cloze	Low Cloze
Karen is **making** a special dinner because it is Tom's birthday today. She is **serving** red wine in crystal glasses and **setting** the table with her finest china. On the stove, she is **grilling** up some steak. In the oven, a chocolate cake is	Baking	Rising
Mrs. Jones 5th grade class is **putting** on a play today. Off stage, some actors are **changing** costumes and **rehearsing** their lines. In front of the stage, the musicians are **playing** their instruments. In the middle of the stage, some ballerinas are	Dancing	Spinning
Jim is **learning** how to live a healthy lifestyle because he wants to lose weight. He has been **making** good progress because he has been **lifting** weights at the gym and **jogging** on the treadmill. Because he hasn't eaten since breakfast, he is hungry and his stomach is	Growling	Grumbling
The Jones family is **preparing** to host a big party this weekend. Mrs. Jones is **doing** the laundry and **folding**	Sweeping	Dusting

clothes. Her son is **sitting** on the couch, but in the kitchen with a broom, her daughter is		
Toby is **staying** at home with the family's two children. He is **operating** a home business and he spends his afternoons **trying** to contact potential customers. Sometimes he has to quit **making** phone calls because the baby wakes up and starts	Crying	Fussing
Some children are outside **making** a snowman. Lily is **putting** a hat on the snowman, and Tom is **placing** sticks for arms. Tomorrow unfortunately, the sun will be **shining** and the temperature will be hot. It's likely that their snowman will start	Melting	Thawing
Jim is **staying** in bed today because his throat has been **hurting** him. He is **taking** his temperature and has been **sucking** on a lozenge because this throat is scratchy and he can't stop	Coughing	Hacking

References

Abramowicz, L. (2007). Sociolingusitics meets exemplar theory: Frequency and recency effects in (ing). *University of Pennsylvania Working Papers in Linguistics, 13*(2), 27-37.

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology, 91*(4), 652-661.

Babel, M. (2010). Dialect Convergence and Divergence in New Zealand English. *Language In Society, 39*, 437-456.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1-29.

Bailey, G. (1997). When Did Southern American English Begin? In E. Schneider (Ed.), *Englishes around the World: Studies in Honour of Manfred Gorlach* (pp. 255-276). Amsterdam: Benjamins.

Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: evidence from ERPs. *Neuroreport, 15*(3), 545-548.

Baugh, J. (2000). Racial identity by speech. *American Speech, 75*, 362-364.

Bell, A. (1984). Language style as audience design. *Language in Society, 13*, 145-204.

Bender, E. (2001). *Syntactic Variation and Linguistic Competence: The Case of AAVE Copula Absence. Stanford University dissertation.*, Stanford University.

Biber, D. S., Johansson, G., Leech, S., Conrad, E., & Finegan., E. (1999). *The Longman grammar of spoken and written English*. London: Longman.

Blaauw, E. (1994). The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech. *Speech Communication, 14*(4), 359-375.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341-345.

Brown, C., & Hagoort, P. (1993). The processing nature of the N400: evidence from masked priming. *Journal of Cognitive Neuroscience, 5*, 34-44.

Brown, V. (1991). Evolution of the merger of /I/ and /E/ before nasals in Tennessee. *American Speech 66*, 303-315.

Bucholtz, M. (2002). From 'Sex Difference' to Gender Variation in Sociolinguistics. *University of Pennsylvania Working Papers in Linguistics, 8*(3), 33-45.

Buchwald, J. S., Guthrie, D., Schwafel, J., Erwin, R. J., & Van Lancker, D. (1994). Influences of language structure on brain-behavior development. *Brain and Language, 46*, 607-619.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change, 14*, 261-290.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory and Language, 56*, 103-128.

Campbell-Kibler, K. (2006). *Listeners perception of sociolinguistic variables: The case of (ING)*. Stanford University.

Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech, 82*(1), 32-64.

Campbell-Kibler, K. (2009). The nature of sociolinguistic perception. *Language Variation and Change, 21*, 135-156.

Campbell-Kibler, K. (2010a). The effect of speaker information on attitudes toward (ING). *Journal of Language and Social Psychology, 29*(2), 214-223.

Campbell-Kibler, K. (2010b). New directions in sociolinguistic cognition. *University of Pennsylvania Working Papers in Linguistics, 15*(2), 31-39.

Campbell-Kibler, K. (2012). The Implicit Association Test and sociolinguistic meaning. *Lingua, 122*(7), 753-763.

Casasanto, L. S. (2010). What do listeners know about sociolinguistic variation? *University of Pennsylvania Working Papers in Linguistics, 15*, 40-49.

Chambers, J. K. (2009). *Sociolinguistic theory : linguistic variation and its social significance*. Oxford: Blackwell.

Cherry, C. (1953). Some experiments on the recognition of speech, with one and with two ears. *J. Acoust. Soc. Am., 25*(5), 975-979.

Clopper, C. G. (2008). Auditory free classification: Methods and analysis. *Behavior Research Methods, 40*(2), 575-581.

Clopper, C. G. (2010). Phonetic Detail, Linguistic Experience, and the Classification of Regional Language Varieties in the United States. In D. R. Preston & N. A. Niedzielski (Eds.), *A Reader in Sociophonetics* (Vol. 219, pp. 203-222). New York: DE GRUYTER.

Clopper, C. G., & Pisoni, D. B. (2004a). Effects of Talker Variability on Perceptual Learning of Dialects. *Language and Speech, 47*(3), 207-238.

Clopper, C. G., & Pisoni, D. B. (2004b). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change, 16*, 31-48.

Clopper, C. G., & Pisoni, D. B. (2005). Perception of Dialect Variation. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 313-337). Malden: Blackwell.

Clopper, C. G., & Pisoni, D. B. (2007). Free classification of regional dialects of American English. *Journal of Phonetics, 35*(3), 421-438.

Cofer, T. (1972). *Linguistic Variability in a Philadelphia Community.*, University of Pennsylvania, Philadelphia.

Conrey, B., Potts, G. F., & Niedzielski, N. A. (2005). Effects of dialect on merger perception: ERP and behavioral correlates. *Brain and Language, 95*, 435-449.

Cutler, A., & Clifton, C. (1999). Comprehending spoken language: a blueprint of the listener. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 123-166). Oxford: Oxford University Press.

Dahan, D., Drucker, S. J., & Scarborough, R. A. (2008). Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition, 108*(3), 710-718.

D'Alessandro, L., & Norwich, K. (2009). Loudness adaption measured by the simultaneous dichotic loudness balance technique differs between genders. *Hearing Research, 247*, 122-127.

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (1999). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology, 36*, 316-328.

De Goede, D., Van Alphen, P., Mulder, E., Blokland, Y., Kerstholt, J., & Van Berkum, J. (2009). *The effect of mood on anticipation in language comprehension: An ERP study*. Paper presented at the 22nd Annual meeting of the CUNY Conference on Human Sentence Processing.

Delong, K. A., Urbach, T., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nat Neurosci, 8*(8), 1117-1121.

Deuchar, M. (1988). A pragmatic account of women's use of standard speech. In J. Coates & D. Cameron (Eds.), *Women in their speech communities* (pp. 27-32). London: Longman.

Dinkin, A. (2004). *The effect of word frequency on F2 of short vowels in American English.* University of Pennsylvania.

Douglas-Cowie, E. (1978). Linguistic code-switching in a Northern Irish village: Social interaction and social ambition. In P. Trudgill (Ed.), *Sociolinguistic Patterns in British English* (pp. 37-51). London: Edward Arnold.

Dreisbach, L., Kramer, S., Cobos, S., & Cowart, K. (2007). Racial and gender effects on pure-tone thresholds and distortion-product otoacoustic emissions (DPOAEs) in normal-hearing young adults. *International Journal of Audiology, 46*, 419-426.

Eckert, P. (1989). The whole woman: sex and gender differences in variation. *Language Variation and Change*, 245-267.

Eckert, P. (2008). Where do ethnolects stop? *International Journal of Bilingualism, 12*(1-2), 25-42.

Elliott, C. (1971). Noise tolerance and extraversion in children. *British Journal of Psychology, 62*, 375-380.

Ervin-Tripp, S. (1973). Children's sociolinguistic competence and dialect diversity. In S. Ervin-Tripp (Ed.), *Language acquisition and communicative choice: Essays by Susan M. Ervin-Tripp* (pp. 262-301). Stanford: Stanford University Press.

Fasold, R. (1990). *The Sociolinguistics of Language*. Cambridge: Basil Blackwell.

Fischer, J. (1958). Social influences on the choice of a linguistic variant. *Word*, 47–56.

Forster, K. (1976). Accessing the mental lexicon. In R. Wales & E. Walker (Eds.), *New Approaches to Language Mechanisms*. Amsterdam: North-Holland.

Fought, C. (1999). A majority sound change in a minority community: /u/-Fronting in Chicano English. *Journal of Sociolinguistics, 3*, 5-23.

Fridland, V., Bartlett, K., & Kreuz, R. (2004). Do you hear what I hear? Experimental measurement of the perceptual salience of acoustically manipulated vowel variants by Southern speakers in Memphis, TN. *Language Variation and Change, 16*, 1-16.

Gardner, R. G., & Lambert, W. (1972). *Attitudes and motivation in second-language learning*. Rowley Massachusetts: Newbury House Publishers.

Garrett, P., Coupland, N., & Williams, A. (2003). *Investigating language attitudes: Social meanings of dialect, ethnicity, and performance*. Cardiff: University of Wales Press.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251-279.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*(1), 4-27.

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*(2), 197-216.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

Greenwald, A. G., Rudman, L. A., Nosek, B. A., Banaji, M. R., Farnham, S. D., & Mellot, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*(1), 3-25.

Gumperz, J. (1968). The speech community *International encyclopedia of the social sciences*: Macmillan.

Guy, G. R. (1980). Variation in the group and the individual: The case of final stop deletion. In W. Labov (Ed.), *Locating Language in Time and Space* (pp. 1-36): Academic Press.

Guy, G. (2007). Variation and phonological theory. In R. Bayley & C. Lucas (Eds.), *Sociolinguistic variation: Theories, methods, and applications*. Cambridge: Cambridge University Press.

Hagoort, P., & Brown, C. M. (1994). Brain responses to lexical ambiguity resolution and parsing. In C. Clifton, L. Frazier & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale: Erlbaum.

Hagoort, P., Hald, L., Bastiaanesen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304*, 438-441.

Hagoort, P., & Van Berkum, J. (2007). Beyond the sentence given. *Philosophical Transactions of the Royal Society, 362*, 801-811.

Hay, J., Nolan, A., & Drager, K. (2006). From Fush to Feesh: Exemplar Priming in Speech Perception. *The Linguistic Review, 23*, 351-379.

Hazen, K. (2008). (ING): A vernacular baseline for English in Appalachia. *American Speech, 83*(2), 116-140.

Henerson, M., Morris, L. & Fitz-Gibbon, C. (1987). How to Measure Attitudes. USA:Sage Publications.

Henrich, J., Heine, S., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61-68.

Hinton, L., Moonwoman, B., Bremner, S., Luthin, H., Van Clay, M., Lerner, J., et al. (1987). *It's not just valley girls: A study of Californian English.* Paper presented at the Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society, Berkeley, CA.

Holcomb, P., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19*(4), 286-300.

Hopper, J. (1976). Word frequency in lexical diffusion and the source of morphological change. In W. Christie (Ed.), *Current Progress in Historical Linguistics*. Amsterdam: North Holland.

Horvath, B. (1985). *Variation in Australian English: The sociolects of Sydney*. New York: Cambridge University Press.

Houston, A. (1985). *Continuity and change in English morphology: the variable (ING).* University of Pennsylvania.

Houston, A. (1991). A grammatical continuum for (ING). In P. Trudgill & J. K. Chambers (Eds.), *Dialects of English: Studies in grammatical variation*. London: Longman.

Jackendoff, R. (2007). *Language, Consciousness, Culture: Essays on Mental Structure*. Cambridge: MIT Press.

Johnson, C. (2006). *Speakin' Tejano: The (ING) variable of Mexican American English of South Texas.* University of Texas, San Antonio.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics, 34*, 485-499.

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

Kinchla, R. A. (1992). Attention. *Annual Review of Psychology, 43*(1), 711-742.

Knops, U., & van Hout, R. (1988). Language attitudes in the Dutch language area: an introduction. In R. van Hout & U. Knops (Eds.), *Language attitudes in the Dutch language area*. Dordrecht: Foris.

Koops, C., Gentry, E., & Pantos, A. (2008). The effect of perceived speaker age on the perception of PIN and PEN vowels in Houston, Texas. *University of Pennsylvania Working Papers in Linguistics, 14*(2), 91-101.

Kutas, M., & Delong, K. A. (2008). A sampler of event-related brain potential (ERP) analyses of language processing. In Z. Breznitz (Ed.), *Brain research in language* (pp. 153-186). New York: Springer.

Kutas, M., & Federmeier, K. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cognitive Science, 4*, 462-470.

Kutas, M., & Federmeier, K. (2011). Thirty Years and Counting: Finding Meaning in the N400 Component of the Event-Related Brain Potential (ERP). *Annual Review of Psychology, 62*(1), 621-647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*, 203-205.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*(5947), 161-163.

Laan, G. P. M. (1997). The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style. *Speech Communication, 22*(1), 43-65.

Labov, W. (1966). *The social stratification of English in New York City*. Washington DC: Center for Applied Linguistics.

Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1984). *The intersection of sex and social factors in the course of language change*. Paper presented at NWAVE.

Labov, W. (1989). The child as linguistic historian. *Language Variation and Change*, 85-97.

Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change, 2*, 205-254.

Labov, W. (2001a). The anatomy of style-shifting. In P. Eckert & J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 85-108). Cambridge: Cambridge University Press.

Labov, W. (2001b). *Principles of linguistic change, Volume 2: Social factors*. Malden: Blackwell Publishers.

Labov, W. (2008). *The cognitive status of sociolinguistic variables. Cognitive capacities of the sociolinguistic monitor*. Paper presented at the Sociolinguistics Symposium 17: Micro and Macro Connections.

Labov, W., Ash, S., Baranowski, M., Nagy, N., Ravindranath, M., & Weldon, T. (2006). Listeners' sensitivity to the frequency of sociolinguistic variables. *Penn Working Papers in Linguistics, 12*(2), 105-129.

Labov, W., Ash, S., & Boberg, C. (2006). *The Atlas of North American English*. Berlin: Mouton-de Gruyter.

Labov, W., Ash, S., Ravindranath, M., Weldon, T., Baranowski, M., & Nagy, N. (2011). Properties of the sociolinguistic monitor. *Journal of Sociolinguistics, 15*(4), 431-463.

Labov, W., Karan, M., & Miller, C. (1991). Near-mergers and the suspension of phonemic contrast. *Language Variation and Change, 3*, 33-74.

Lambert, W., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology, 60*.

Lattner, S., & Friederici, A. (2003). Talker's voice and gender stereotype in human auditory sentence processing - evidence from event-related brain potentials. *Neuroscience Letters, 339*, 191-194.

Lau, E., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews: Neuroscience, 9*, 920-933.

Lee, R. R. (1971). Dialect Perception: a Critical Review and Re-Evaluation. *Quarterly Journal of Speech, 57*(410-417).

Leinonen, A., Grönholm-Nyman, P., Järvenpää, M., Söderholm, C., Lappi, O., Laine, M., & Krause, C. M. (2009). Neurocognitive processing of auditorily and visually presented inflected words and pseudowords: Evidence from a morphologically rich language. *Brain Research, 1275*, 54-66.

Levine, L., & Crockett, H. (1966). Speech variation in a piedmont community: Post-vocalic /r/. In S. Lieberson (Ed.), *Explorations in Sociolinguistics* (pp. 76-98). The Hague: Mouton and Co.

Loudermilk, B. (in press). Psycholinguistic Approaches. In R. Bayley, R. Cameron & C. Lucas (Eds.), *The Oxford Handbook of Sociolinguistics*: Oxford University Press.

Loudermilk, B., Gutierrez, E., & Corina, D. (submitted). What are you talkin' about? Perception of sociolinguistic variation: Evidence from event-related potentials

Luck, S. (Ed.). (2005). *An introduction to the event-related potential technique*. Cambridge: MIT Press.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The Weckud Wetch of the Wast: Lexical Adaptation to a Novel Accent. *Cognitive Science, 32*(3), 543-562.

Martin-Loeches, M., Schacht, A., Casado, P., Hohlfeld, A., Rahman, R. A., & Sommer, W. (2009). Rules and heuristics during sentence comprehension: Evidence from a dual-task brain potential study. *Journal of Cognitive Neuroscience, 21*(7), 1365-1379.

McGuinness, D. (1974). Equating individual differences for auditory input. *Psychophysiology, 11*, 115-120.

Mencken, H. L. (1963). *The American Language* (4th ed.). New York: Alfred A. Knopf.

Morton, J. (1970). A functional model of human memory. In D. Norman (Ed.), *Models of human memory*. New York: Academic Press.

Naatanen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., & Iivonen, A. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*, 432-434.

Niedzielski, N. A. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology, 18*(1), 62-85.

Nieuwland, M. S., & Van Berkum, J. (2006). When peanuts fall in love: N400 evidence of the power of discourse. *Journal of Cognitive Neuroscience, 18*(7), 1098-1111.

Nygaard, L., & Pisoni, D. B. (1995). Speech Perception: New Directions in Research and Theory. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of Perception and Cognition*. San Diego: Academic Press.

Osterhout, L., & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family. *Language and Cognitive Processes, 14*(1), 1-14.

Osterhout, L., & Holcomb, P. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785-806.

Ozyurek, A., Willems, R. M., Kita, S., & Hagoort, P. (2007). On-line integration of semantic information from speech and gesture: Insights from event-related potentials. *Journal of Cognitive Neuroscience, 19*(4), 605-616.

Pantos, A. (2010). *Measuring implicit and explicit attitudes toward foreign-accented speech.* Rice University.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure*. Philadelphia: John Benjamins.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition, 95*, 201-236.

Plichta, B., & Preston, D. R. (2005). The /ay/s have it: The perception of /ay/ as a North-South stereotype in US english. *Acta Linguistica Hafniensia*, 243-285.

Plichta, B., & Rakerd, B. (2010). Perceptions of /a/ fronting Across Two Michigan Dialects. In D. R. Preston & N. A. Niedzielski (Eds.), *A Reader in Sociophonetics* (Vol. 219, pp. 223-240). New York: DE GRUYTER.

Preston, D. R. (1993). Folk Dialectology. In D. R. Preston (Ed.), *American Dialect Research* (pp. 333-377). Amsterdam: Benjamins.

Purnell, T., Idsardi, W. J., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology, 18*, 10-30.

Purnell, T., Raimy, E., & Salmons, J. (2009). Defining dialect, perceiving dialect, and new dialect formation: Sarah Palin's speech. *Journal of English Linguistics, 37*(4), 331-255.

Radford, A. (1997). *Syntactic theory and the structure of English: A minimalist approach*. United Kingdom: Cambridge University Press.

Regel, S., Gunter, T. C., & Friederici, A. D. Isn't it ironic? an electrophysiological exploration of figurative language processing. *J. Cognitive Neuroscience, 23*(2), 277-293.

Reid, E. (1978). Social and stylistic variation in the speech of children: Some evidence from Edinburgh. In Douglas-Cowie (Ed.), (pp. 158–171).

Rickford, J. R. (1999). *African American Vernacular English: Features, Evolution,Educational Implications*: Blackwell.

Rivera-Gaxiola, M., Osibra, G., Johnson, M. H., & Karmiloff-Smith, A. (2000). Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behavioral Brain Research, 111*, 13-23.

Roberts, J. L. (1994). *Acquisition of Variable Rules: (-t,d) Deletion and (ing) Production in Preschool Children.* University of Pennsylvania, Philadelphia.

Rogers, D., Harkrider, A., Burchfield, S., & Nabelek, A. (2003). The influence of listener's gender on the acceptance of background noise. *Journal of the American Academy of Audiology, 14*, 372-382.

Russell, J. (1982). Networks and sociolinguistic variation in an African urban setting. In S. Romaine (Ed.), *Sociolinguistic Variation in Speech Communities* (pp. 125-140). London: Edward Arnold.

Sagi, E., D'Alessandro, L., & Norwich, K. (2007). Identification variability as a measure of loudness: an application to gender differences. *Canadian Journal of Experimental Psychology, 61*, 64-70.

Savin, H. (1963). Word-Frequency Effect and Errors in the Perception of Speech. *Journal of the Acoustical Society of America 35*, 200-206.

Schilling-Estes, N. (2002). American English Social Dialect Variation and Gender. *Journal of English Linguistics, 30*(2), 122-137.

Shahnaz, N. (2008). Transient evoked otoacoustic emissions (TEOAEs) in Caucasian and Chinese young adults. *International Journal of Audiology, 47*, 76-83.

Shopen, T. (1978). Research on the variable (ING) in Canberra, Australia. *Talanya, 5*, 42-52.

Shuy, R., Wolfram, W., & Riley, W. (1967). *Linguistic correlates of social stratification in Detroit speech*: USOE Final Report.

Smith, M. E., & Halgren, E. (1987). Event-related potentials during lexical decision: Effects of repetition, word frequency, pronounceability, and concreteness. *Electroencephalogr. Clin. Neurophysiol. Suppl. , 40*, 417-421.

Squires, L. M. (2011). *Sociolinguistic priming and the perception of agreement variation: Testing predictions of exemplar-theoretic grammar.*

Staum, L. (2008). *Experimental investigation of sociolinguistic knowledge.* Stanford University.

St. George, M., Mannes, S., & Hoffman, J. E. (1994). Global semantic expectancy and language comprehension. *Journal of Cognitive Neuroscience, 6*, 70-83.

Strand, E. A., & Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. In D. Gibbon (Ed.), *Natural language processing and speech technology*. Berlin: Mouton.

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final /t. *Journal of Memory and Language, 52*(3), 322-338.

Sumner, M., & Samuel, A. G. (2009). The effect of experience on the perception and representation of dialect variants. *Journal of Memory and Language, 60*(4), 487-501.

Tanenhaus, M. K., & Brown-Schmidt, S. (2008). Language processing in the natural world. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*(1493), 1105-1122.

Thomas, E. (2002). Sociophonetic applications of speech perception experiments. *American Speech, 77*(2), 115-147.

Trudgill, P. (1972a). Sex, covert prestige and linguistic change in the urban British English of Norwich *Language In Society, 1*, 179-196.

Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

van Alphen, P., & van Berkum, J. (2009). Is There Pain in Champagne? Semantic Involvement of Words within Words during Sense-making. *Journal of Cognitive Neuroscience, 22*(11), 2618-2262.

Van Berkum, J., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: evidence from the N400. *Journal of Cognitive Neuroscience, 11*(6), 657-671.

Van Berkum, J., Holleman, B. C., Murre, J. M. C., Nieuwland, M. S., & Otten, M. (2007). *So how do you feel about this? An ERP study on opinion poll comprehension*. Paper presented at the Annual Meeting of the Cognitive Neuroscience Society.

Van Berkum, J., van den Brink, D., Tesink, C., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20*(4), 580-591.

Van Berkum, J., Zwitserlook, P., Hagoort, P., & Brown, C. M. (2003). When and how listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research, 17*, 701-718.

van den Brink, D. l., Van Berkum, J. J. A., Bastiaansen, M. C. M., Tesink, C. M. J. Y., Kos, M., Buitelaar, J. K., & Hagoort, P. (2010). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social Cognitive and Affective Neuroscience*.

Wald, B., & Shopen, T. (1985). A researcher's guide to the sociolinguistic variable (ING). In V. P. Clark, P. A. Eschholz & A. G. Rosa (Eds.), *Language: Introductory readings* (4th ed., pp. 515-541). New York: St. Martin's Press.

Weinrich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for Historical Linguistics.* Austin: University of Texas Press.

Willems, R. M., Ozyurek, A., & Hagoort, P. (2008). Seeing and hearing meaning: ERP and fMRI evidence of word versus picture integration into a sentence context. *Journal of Cognitive Neuroscience, 20*(7), 1235-1249.

Williams, J., & Themanson, J. (2011). Neural correlates of the implicit association test: evidence for semantic and emotional processing. *Social Cognitive and Affective Neuroscience, 6*(4), 468-476.

Wolfram, W., & Christian, D. (1976). *Appalachian Speech*. Arlington: Center for Applied Linguistics.

Wolfram, W., & Schilling-Estes, N. (1998). Gender and language variation. In B. Publishing (Ed.), *American English*.

Woods, H. (1978). *A socio-dialectal survey of the English spoken in Ottawa: A study of sociolinguistics and stylistic variation.* University of Ottawa.