



University of Tennessee Health Science Center
UTHSC Digital Commons

Theses and Dissertations (ETD)

College of Graduate Health Sciences

12-2012

N200 and P300 Evoked by Stimuli Straddling Category Boundary in Lexical Context

Ji Young Lee

University of Tennessee Health Science Center

Follow this and additional works at: <https://dc.uthsc.edu/dissertations>

 Part of the [Medical Sciences Commons](#)

Recommended Citation

Lee, Ji Young, "N200 and P300 Evoked by Stimuli Straddling Category Boundary in Lexical Context" (2012). *Theses and Dissertations (ETD)*. Paper 147. <http://dx.doi.org/10.21007/etd.cghs.2012.0175>.

This Dissertation is brought to you for free and open access by the College of Graduate Health Sciences at UTHSC Digital Commons. It has been accepted for inclusion in Theses and Dissertations (ETD) by an authorized administrator of UTHSC Digital Commons. For more information, please contact jwelch30@uthsc.edu.

N200 and P300 Evoked by Stimuli Straddling Category Boundary in Lexical Context

Document Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Program

Audiology and Speech Pathology

Research Advisor

Ashley W. Harkrider, Ph.D.

Committee

Mark S. Hedrick, Ph.D. Sky Huck, Ph.D. Deborah von Hapsburg, Ph.D.

DOI

10.21007/etd.cghs.2012.0175

**N200 and P300 Evoked by Stimuli Straddling Category Boundary
in Lexical Context**

A Dissertation
Presented for
The Graduate Studies Council
The University of Tennessee
Health Science Center

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy
From The University of Tennessee

By
Ji Young Lee
December 2012

Copyright © 2012 Ji Young Lee.
All rights reserved.

DEDICATION

To God and my family.

ACKNOWLEDGEMENTS

First of all, I thank God for the journey I have been ever through. I wholeheartedly appreciate my family's love and support. They are the best gift in my life. I would like to express my deep respect and gratitude to my advisor, Dr. Ashley Harkrider for her mentorship and care. She is and will be my role model in my future career. I thank my committee members, Dr. Mark Hedrick, Dr. Deborah von Hapsburg, and Dr. Sky Huck. Their invaluable feedback and encouragement always cheer me up so that I can move on. Also, I thank Max, Ron, Pam, and Karen for their technical and administrative support. It was very fortunate to work with them. Lastly, I thank Gong, Yoon Hee, Sun Mi, Nam Jun, Dania, Matt, Ruth, and Barbara. I am so blessed to have them as friends.

ABSTRACT

Background and Objectives: Event-related potentials (ERPs) like N200 and P300 have been reported to reflect the categorical perception of speech. The purpose of the present study is to explore whether these ERP components reflect the influence of lexical context on categorical perception. Findings may provide evidence for bottom-up or top-down processing of speech.

Methods: On a seven step series of the /bi/-/pi/ continuum, a two-forced choice labeling test was administered in two conditions: /bi/ context (e.g. bee sting) and /pi/ context (e.g. pea soup). From the labeling results, Stimulus 1 which is a prototypical /bi/ was selected for standard stimuli, and Stimulus 4 which showed the greatest effect of context of the between-category was selected as the deviant in an active oddball paradigm commonly used to obtain N200 and P300 ERPs. After subjects finished a two-forced choice labeling test, they participated in electrophysiological testing while simultaneously pressing a response button when they heard the deviant stimuli. A total of 450 stimuli composed of 369 standard stimuli (81%) and 81 deviant stimuli (19%) were presented in an active oddball paradigm for /bi/ and /pi/ context word conditions, respectively. ERP responses were measured using 9 electrodes from 21 normal hearing adults. Electrophysiological data (amplitude and latency of N200 and P300) and behavioral data (labeling, discrimination response accuracy, discrimination reaction time) were analyzed.

Results: (1) The amplitude and latency of N200 and P300 did not represent the change in categorical perception in the presence of lexical context that was demonstrated in the labeling task, (2) N200 amplitude was the largest over the frontal region while P300 was the largest over the parietal region, (3) discrimination reaction time was faster in the /pi/ context condition than /bi/ context condition while response accuracy did not differ with context, and (4) there was no correlation between N200/P300 and behavioral data.

Conclusion: N200 and P300 do not reflect the lexical effect on pre-lexical processing of categorical speech stimuli. Findings suggest that lexical context does not affect electrophysiological measures of pre-lexical speech processing (e.g., N200 or P300), supporting the autonomous (bottom-up) model that speech perception influenced by lexical context (as demonstrated behaviorally) occurs on the post-lexical level.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION	1
Categorical Perception	1
Event-related Potentials	2
N1 and P2	2
N200	3
P300	4
Event-related Potentials on Categorical Perception	4
Rationale and Research Questions	5
CHAPTER 2. LITERATURE REVIEW	7
Models of Speech Perception: Autonomous vs Interactive	7
Categorical Perception	8
Categorical Perception of Speech	8
Categorical Perception of Speech in Lexical Context	9
Event-related Potentials on Categorical Perception of Speech	10
Research Questions and Hypotheses	13
CHAPTER 3. METHODOLOGY	15
Subjects	15
Stimuli	15
Procedure	16
Labeling Test	16
Electrophysiological Test and Discrimination Test	16
Data Analysis	18
Labeling Test	18
Electrophysiological Test	18
Discrimination Test	19
Correlation between Electrophysiological and Behavioral Tests	19
CHAPTER 4. RESULTS	20
Labeling Data	20
Electrophysiological Data	20
N200	20
P300	20
N200-P300 Complex	25
Discrimination Data	25
Correlation between EEG and Behavioral Data	25
CHAPTER 5. DISCUSSION	30
Interpretation of the Current Results	30
Limitations and Suggestion for Future Research	32
Conclusion	33

LIST OF REFERENCES.....	35
VITA.....	42

LIST OF TABLES

Table 1-1.	Stimuli for two conditions in an oddball paradigm.	6
Table 4-1.	Post-hoc tests for the interaction of lexical context and VOT.	21
Table 4-2.	N200 amplitude and latency by scalp distribution.	24
Table 4-3.	P300 amplitude by scalp distribution.	26
Table 4-4.	N2-P3 complex by scalp distribution.	26
Table 4-5.	Discrimination response accuracy and reaction times by subject.	27
Table 4-6.	Bivariate correlation in /bi/ and /pi/ context conditions.	29

LIST OF FIGURES

Figure 3-1. /bi/ response in /bi/ and /pi/ contexts by VOT.	17
Figure 4-1. Grand average waveforms to the standard stimuli (dashed) and to the deviant stimuli (solid) in /bi/ context.	22
Figure 4-2. Grand average waveforms to the standard stimuli (dashed) and to the deviant stimuli (solid) in /pi/ context.	23
Figure 4-3. Grand average deviant waveforms at Fz magnified in /bi/ context (dashed) and /pi/ context (solid).	24
Figure 4-4. Lexical effect on behavioral responses.	28

LIST OF ABBREVIATIONS

ANSI	American National Standard Institute
dB SPL	Decibel Sound Pressure Level
EEG	Electroencephalography
ERP	Event-related Potential
Hz	Hertz
MMN	Mismatch Negativity
PMN	Phonological Mapping Negativity
VOT	Voice Onset Time

CHAPTER 1. INTRODUCTION

A speech sound contains complex, dynamic, and rapidly changing acoustic information and generally is realized in the form of the continuous stream of speech. Thus, a high-level cognitive function is required to extract crucial acoustic cues from the continuous acoustic features. Speech perception occurs on the mapped representation of acoustic input to the phonological categories.

There have been numerous studies to examine how human speech sounds are processed and finally integrated into language. Of them, phoneme perception has been of interest as a basic step. Over the past half century, it has been a central topic in speech perception how auditory sensory information is translated into phonetic representation. Categorical perception of speech has been studied widely to examine the phonetic process since Liberman, Harris, Hoffman, and Griffith (1957) first presented their model.

Categorical Perception

Categorical perception is a well known pattern of information processing in human cognitive activity. Categorical perception occurs when a certain sensory input is conveyed to the brain and processed into a certain abstract category rather than the physical reality itself. The human brain has adapted to process existing innumerable information efficiently by sorting it into fixed discrete categories. In the presence of a great amount of continuous sensory input, stimuli with the same physical differences can be perceived into the same or different categories depending on the category system represented in the brain (Harnad, 1987). Categorical perception helps the brain process the information economically by ignoring unnecessary information and picking up crucial information only. Also, stimuli can be discriminated better when they belong to different categories than the same categories, though they have the same physical differences (Liberman, Harris, Hoffman, & Griffith, 1957).

The way lexical context affects categorical perception has been studied. Humans can recognize words made of sequences of speech sounds even in difficult listening conditions. We often do not perceive the exact sensory information of a speech sound in an everyday life situation due to the speaker (unintelligible speech), listener (poor hearing ability), or environmental factors (noise or reverberation). However, we learn to overcome uncertainty when acoustic information is absent or ambiguous, in part by using semantic cues. Particularly, the ‘Ganong effect’, which refers to a shift in the category boundary depending on the lexical status of the context, has been widely studied to show the strong influence of lexical context on phonetic categorization. For example, an ambiguous segment between /g/ and /k/ may be perceived as /g/ in a context that forms the word ‘gift’, or as ‘k’ in a context that forms the word ‘kiss’ (Ganong, 1980; Pitt & Samuel, 1993).

Event-related Potentials

Electroencephalography (EEG) has been a useful tool to study speech and language processing due to its objective, non-invasive assessment, and high temporal resolution. Electrophysiological measurements provide objective information about normal and abnormal auditory function in the field of audiology.

Of them, to date, Event-related potentials (ERPs) frequently have been used to examine brain function during cognitive tasks. ERPs refer to the electrical brain activities which are temporally synchronized to a given event, such as a sound, phoneme, word, or sentence. Scalp-recorded ERPs are time-locked responses of electrical activity representing a specific aspect of cortical processing (Patel & Azzam, 2005). They are primarily summed post-synaptic potentials of synchronously activated pyramidal cells in the neocortex (Kutas & Federmeier, 2000). They provide information regarding how, when, and (very generally) where speech processing occurs during a given task (Hillyard & Picton, 1987). Typically, amplitude, the strength of the responses, is interpreted as reflecting the size and synchronicity of the neural population involved in a given processing task. Latency, the time period when the responses occur, is thought to imply the amount of time that it takes for a specific stimulus to reach the central auditory nervous system.

ERPs to speech sounds may be of great value to examine biological and neural processes underlying speech and language processing. They also can be clinically useful in assessing speech and language performance, planning rehabilitation, and monitoring progress (Martin, Tremblay, & Korczak, 2008). N200 and P300 are not restricted to auditory tasks and appear to be associated closely with the cognitive processes of perception, selective attention, and conscious discrimination in humans (Patel & Azzam, 2005). Thus, these ERP components seem to be optimal to study how lexical context affects the categorical perception of speech sounds.

N1 and P2

N1 and P2 have been known to reflect preattentive processing of sound such as speech segmentation and detection. They are thought to be sensitive to the acoustic properties of sound as obligatory sensory-evoked potentials (Kutas, Delong, & Kiang, 2011). While N1 is evoked primarily over the fronto-central region of the scalp, the scalp distribution of P2 is more diffusely distributed across fronto-central and parieto-occipital regions (e.g., Näätänen and Picton, 1987; Ross and Tremblay, 2009). In the context of ERP studies, N1 and P2 are often used in identifying the later components (e.g., N200, P300) for analysis purposes.

N200

N200, which is evoked at around 180 to 325 ms following the presentation of a stimulus, is the negativity caused by a deviation from preceding stimuli. The N200 typically is evoked before a motor response, suggesting its link to the cognitive processes of stimulus labeling and distinction. The N200 is better observed in the difference waveform obtained by subtracting the waveform to the standard stimuli from the waveform to the target stimuli (Wang, Miyazato, Randall, Hokama, Hiramatsu, & Ogura, 2003). Though there have been some inconsistent findings about N200 in auditory and visual modalities, N200 is thought to reflect cognitive processes beyond the detection of stimulus mismatch or attention, such as monitoring, and regulation, feedback of information (Jonathan, Folstein, & Van Petten, 2008).

N200 was thought to have two components, N2a and N2b, until the early 1980s. Since the N2c was newly proposed (Ritter, Simson, Vaughan, & Friedman, 1979; Ritter, Simson, Vaughan, & Macht, 1982), three subcomponents of N200 evoked by auditory stimuli have been characterized. The N2a is elicited maximally from the anterior scalp distribution around 100~250 ms after stimulus onset. It is evoked whether subjects pay attention to or ignore the stimulus (Näätänen et al., 2007; Patel & Azzam, 2005). N2a is known to reflect the automatic processing of the disparity between a mismatched stimulus and a sensory memory. N2a commonly is referred to as MMN when evoked by auditory stimuli and it has its visual equivalent (vMMN) and magnetoencephalographic equivalent (MMNm) (Hari et al., 1984). Näätänen et al. (1978) first isolated MMN from N200 on the standard-deviant difference waveform. N2a is known to reflect automatic change detection mechanisms based on memory traces, thus MMN is elicited when the auditory input and the prediction from a memory trace do not match. N2a is known to be independent of subjects' attention though some studies (Woldorff, Hackley, & Hillyard, 1991; Woldorff, Hillyard, Gallen, Hampson, & Bloom, 1998) show that MMN amplitude decreases under heightened attention.

On the other hand, N2b and N2c are elicited only when attention is required. Specifically, N2b is assumed to reflect the detection of a stimulus mismatch. N2b is elicited along with P3a, forming N2b-P3a complex on a central scalp distribution (somewhat posterior to N1 and MMN) whereas N2c is elicited along with P3b on a fronto-central scalp distribution. Hence, all subcomponents of N200 evoked by auditory stimuli show a frontocentral or central scalp distribution whereas N2c has posterior scalp distribution in the visual modality (Folster & Van Petten, 2008; Patel & Azzam, 2005). The N2a and N2b/P3a complex are thought to index the mismatch detection to trigger the orienting reflex whereas N2c is thought to reflect a subprocess of classification tasks (Folstein & van Petten, 2008). The N2b response has a more posterior scalp distribution compared to N1 or MMN. The scalp distribution of N2c has been reported as superior frontal and temporal regions (Dehaene-Lambertz, 1997), centrality of the frontal and superior temporal cortex (Potts, Dien, Hartry-Speiser, McDougal, & Tucker, 1998), and left temporoparietal region (Celsis, Doyon, Boulanouar, Pastor, Demonet & Nespoulous, 1999). N200 refers to N2c in the present study.

P300

Since it was described first by Sutton et al. (1965), the P300 has been the most frequently studied ERP component, possibly because the response is robust and reliable to obtain in experimental environments (Patel & Azzam, 2005). P300 is a positive-going peak at around 250~500 ms after stimulus onset, and is elicited in the oddball paradigm while listeners attend to the stimuli passively or actively. The resultant parietal P300 is modulated by the listeners' attention status, probability of deviant stimuli, and task difficulty, such that the less demanding the task, the greater the amplitude of the P300 (Polich, 2007). P300 is known to reflect selective attention, broad recognition, and memory-updating processes. P3b in particular is thought to index a working memory trace.

As a subcomponent of P300, the involuntary component of the P300 is called the P3a, which is an earlier positive peak with frontal or fronto-central scalp topography. P3a typically is found in young adults and its amplitude decreases over relatively few trials (Bennington & Polich, 1999). A frontally maximal P3a is known to reflect the orienting of attention. On the other hand, the P3b is elicited by the oddball stimuli when the listener actively attends. It has a longer latency and more posterior (centro-parietal) scalp topography compared to P300a. A parietally maximal P3b is known to reflect the updating of working memory. Generally, when studies in the literature cite P300, they are referring to P3b. The P3b is a voluntary, obligatory, and reliable response with low intra- and inter-subject variability (Jonathan, Folstein, & van Petten, 2008). P300 refers to P3b in the present study.

Event-related Potentials on Categorical Perception

Most studies measuring speech evoked ERPs to categorical perception have used within-category or across-category stimuli. They have examined whether a specific ERP response is sensitive to acoustic versus phonetic change. Of these, most have focused on the MMN response and aimed to see whether MMN reflects the detection of acoustic or phonetic change. The results are still controversial. Some studies show no MMN is elicited to stimuli within the same phonemic category (Aaltonen, Paavilainen, Sams, & Näätänen, 1992). Other studies report that MMN is more dependent on acoustic changes than phonetic changes (Maiste, Wiens, Hunt, Scherg, & Picton, 1995; Sams, Aulanko, Aaltonen, & Näätänen, 1990; Sharma, Kraus, McGee, Carrell, & Nicol, 1993) or MMN reflects both acoustic and phonetic processing (Näätänen et al., 1997; Peltola, Kujala, Tuomainen, Ek, Aaltonen, & Näätänen, 2003; Sharma & Dorman, 1999; Tampas, Harkrider, & Hedrick, 2005; Winkler, Lehtokoski, Alku, Vainio, Czigler, Csepe, et al. 1999).

P300 has been thought to represent conscious discrimination of sound. However, the results of P300 response to speech sound are also controversial. As in MMN, P300 may represent acoustic changes in speech sound (Martin, Tremblay, & Korezak, 2008;

Tampas, Harkrider, & Hedrick, 2005) or phonetic changes (Dalebout & Stack, 1999; Dehaene-Lambertz, 1997; Maiste, Wiens, Hunt, Scherg, & Picton, 1995).

Rationale and Research Questions

In the real world, the same speech sound can be perceived differently due to factors such as phonetic context, lexical context, and/or frequency of use. Therefore, it is crucial to study speech perception not only in isolation but also in a more ecological environment. The Ganong effect shows how categorical perception of speech can be influenced by lexical context using behavioral measures. However, it is not known how the neural responses underlying lexical effect of categorical perception appear nor at what level of the central nervous system. Because the Ganong effect is associated with both speech and language processing, the current study could provide neurocognitive information of speech and language processing and contribute to developing clinical application.

The neurophysiologic process underlying speech perception is still elusive. In spite of numerous auditory evoked ERP studies, N200 and P300 responses to speech sounds have not been widely studied. Most auditory evoked ERP studies have examined N200 and P300 focusing on auditory or speech processing by using nonspeech sounds or meaningless CV syllables in isolation. Moreover, little is known about the neurophysiologic nature and function of N200 possibly because N200 is hard to detect in experimental studies; N200 shows large variability in latency, overlaps with P200, and appears to be associated with both match and mismatch conditions on a given task (Wang, Miyazato, Randall, Hokama, Hiramatsu, & Ogura, 2003). Though N200 and P300 are assumed to reflect high-level cognitive processes required for speech perception, they have been little studied on the lexical level. To examine the lexical effect on categorical perception of speech, however, N200 and P300 may be optimal ERP components because N200 and P300 have been reported to index controlled processing related to the detection of sound change. Thus, these responses may reflect the change of speech perception in word stimuli with lexical context as well.

Specifying the above research topic, between-category stimuli were used because as previous behavioral research shows, the perception of ambiguous sounds is most influenced by lexical context. Also, most speech evoked ERP studies have used within-category stimuli which are very difficult to discriminate or cannot be discriminated, or across-category stimuli which are very easy to discriminate, behaviorally. Little is known about whether ERPs will be elicited to between-category stimuli. Between-category speech stimuli are acoustically different but phonetically transitional sounds. Thus, ERP responses to between-category stimuli could provide additional detailed information about neural models of speech perception.

The objective of the present study is to examine whether ERP components reflect the lexical effect on categorical perception of speech measured behaviorally. To do this, (1) a seven step series of /bi-/pi/ continuum was made by manipulating VOT, (2) a two-

forced choice labeling test on the continuum was administered in two conditions of /bi/ context word (e.g. bee sting) and /pi/ context word (e.g. pea soup), (3) from the labeling test results, Stimulus 1(e.g. /bi/) was selected as standard stimuli and Stimulus 4 was selected as deviant stimuli for an active oddball paradigm, (4) an oddball paradigm was constructed in two conditions of /bi/ context and /pi/ context, respectively, and (5) electrophysiological (N200 and P300) and behavioral tests (labeling, discrimination response accuracy, discrimination reaction time) were analyzed. **Table 1-1** shows the stimuli used for two conditions in an oddball paradigm.

The research questions in the present study are four fold.

1. Do the ERPs (N200 and P300) reflect the lexical effect on categorical perception of speech?
2. How does the scalp distribution of N200 and P300 appear on the given tasks?
3. Do the behavioral data (labeling, discrimination response accuracy, and discrimination reaction times) reflect the lexical effect on categorical perception of speech?
4. Do ERP and behavioral data correlate with each other?

Table 1-1. Stimuli for two conditions in an oddball paradigm.

Lexical context	Standard stimuli	Deviant stimuli
/bi/ context condition	Stimulus 1 -/stiŋ/	Stimulus 4-/stiŋ/
/pi/ context condition	Stimulus 1-/sup/	Stimulus 4-/sup/

CHAPTER 2. LITERATURE REVIEW

Models of Speech Perception: Autonomous vs Interactive

For the past half century, researchers have been trying to delineate how speech sounds are processed, particularly whether the flow of information through the system is strictly bottom-up, or if there is also a top-down component. It has been one of the central debates in the speech perception field whether there is only feedforward from acoustic and phonemic processing to word recognition or there is also feedback from lexical level to sub-lexical level. There are currently both strong proponents for bottom-up processing, in which successively more abstract representations are analyzed, and top-down processing, in which activation of a more abstract representation can affect the activation of smaller units.

In bottom-up or autonomous models, the incoming speech signal is initially processed independent of context in a bottom-up manner. These models maintain that word recognition is solely bottom-up processing, with speech flowing through the prelexical system in one direction (Culter, Mehler, Norris, & Segui, 1987; Massaro, 1989, Norris, 1994), with the influence of lexical and semantic contextual information arising only at a later decision stage. For example, Shortlist (Norris, 1994) and Merge (Norris, McQueen, and Cutler, 2000) models are feedforward, having only bottom-up connections between the phoneme and lexical levels of representation. Although they provide an account for different types of competition, it is proposed that these arise from the influence of lexical knowledge on phonological processing at the decision stage. They argue that information flow from word processing to these earlier stages is not required by the logic of speech recognition, thus it could only have been included in models of speech recognition as an additional component. These models explain listeners' bias to interpret the ambiguous sound in a way that is consistent with a word takes place at a level of processing that occurs after phoneme labeling, and the lexical influence of word knowledge never feeds back to the stage of phoneme labeling.

In a top-down model or interactive model, on the other hand, our low-level perceptual experiences are influenced by top-down feedback from higher levels of processing. Thus, what we hear is strongly affected by what we expect to hear and by the knowledge from our previous experiences (Connine & Clifton, 1987; McClelland & Elman, 1986; Samuel, 1996; Samuel, 2001). According to this model, the contextual effect occurs due to a lexical influence on pre-lexical representation. This interactive view predicts that lexical information actually reaches down and remaps the mental representation of the sound that is already heard. For example, the TRACE model (McClelland & Elman, 1986) claims that reciprocal interaction occurs between lexical and pre-lexical levels permitting a bi-directional flow (feedforward and feedback) of information during word recognition. These models assume that higher-level language processes connected with morphology, syntax, or semantics may interact with basic speech perception processes. Thus, listeners are able to compensate for missing or noise-

masked phonemes using their knowledge about phonemic structure from the acoustic signal.

Autonomous models have feedforward excitatory connections from pre-lexical level to lexical level and a separate phoneme labeling processing that integrates the information from both pre-lexical and lexical levels. Information in the lexical level and phoneme labeling level is processed through inhibitory connections within each level. On the other hand, interactive models have bi-directional excitatory connections from pre-lexical level to lexical level, or from lexical level to pre-lexical level. Information in the lexical level and pre-lexical level is processed through inhibitory connections within each level (McClelland, Mirman, & Holt, 2006).

The models of speech perception could be tested by specifically designed experiments of speech and language processing. The categorical perception is basically involved with pre-lexical level while the lexical effect is associated with lexical level. Thus, ERP data in addition to behavioral data on the categorical perception influenced by lexical effect could show what occurs during speech and language processing in more detail and how it can be interpreted in the models of speech perception.

Categorical Perception

Categorical perception means that a gradual change along a continuum is perceived as an instance of discrete categories, not as continuous change. It is called categorical perception because we perceive an ideal example of a particular category instead of perceiving ambiguous information itself. Therefore, even when the physical stimuli change continuously, people perceive it categorically. This effect has been widely explored on stimuli varying along a continuum, such as speech sound, color, and facial expressions (Bornstein & Korda, 1984; Etcoff & Magee, 1992; Young et al., 1997).

Categorical Perception of Speech

Categorical perception of speech was studied first at the Haskins Laboratories in 1957. Liberman, Harris, Hoffman, and Griffith (1957) constructed a synthetic /ba/-/da/-/ga/ continuum where the syllables were manipulated in the second formant transition. The experiment revealed that categorical perception shows steep labeling functions, and peaks in the discrimination functions appear between stimuli that span the (50%) labeling boundary.

To date, numerous experiments have examined the phenomenon of categorical perception, using a series of consonant-vowel (CV) continuum synthesized to vary an acoustic parameter in equal steps along a continuum. Thus, categorical perception has been established for a lot of dimensions of phonetic contrast such as place of articulation, manner of articulation, and voicing. For example, a continuum of CV syllables varying in the slope of the F2 transition (e.g. /ba/-/da/-/ga/), the bandwidth of the fricative noise (e.g.

/s-/ʃ/ continuum), and the duration of VOT (e.g. /ba-/pa/, /da-/ta/, /ga-/ka/) were presented to listeners for phonemic labeling or discrimination to demonstrate categorical perception. Results indicated that consonants were perceived categorically though fricatives seem to be perceived somewhat less categorically. Generally, categorical perception has two patterns: (1) labeling functions showed abrupt boundaries between phoneme categories; (2) discrimination functions showed discrimination accuracy was low for within-category stimulus but high for between-category stimuli, particularly for the stimuli that straddled the 50% labeling mark.

Categorical Perception of Speech in Lexical Context

The lexical effect on phoneme labeling has been proved in two aspects of labeling shift and the reaction time effect; ambiguous phonemes tend to be perceived as lexically consistent (Ganong, 1980; Pitt & Samuel, 1993) and phonemes are recognized more quickly in words than nonwords (Connine & Clifton, 1987; Cutler et al, 1987).

In particular, the evidence for lexical effects has come from results that phonetic category boundaries were altered by the surrounding context. Ganong (1980) first suggested that phoneme perception can be influenced by linguistic context. He synthesized seven pairs of seven-step synthetic continua, manipulating VOTs from 15 ms to 55 ms. For each pair from the continuum, one endpoint was a word (e.g. ‘dash’, ‘task’) and the other endpoint was a nonword (e.g. ‘tash’, ‘dask’). He suggested lexical status affects phoneme perception based on the finding that an ambiguous consonant in the continuum was identified more often as a word than a nonword. For example, an ambiguous segment between /g/ and /k/ is more likely to be identified as ‘/g/’ if followed by a /g/-primed context (e.g. ‘gift’), or as ‘/k/’ if followed by a /k/-primed context (e.g. ‘kiss’). He argued that this lexical labeling shift is evidence of lexical knowledge feeding back to phoneme processing.

Borsky, Tuller, and Shapiro (1998) investigated how sentence context influences phoneme categorization. They constructed a ten-step continuum of ‘goat’ and ‘coat’ by manipulating VOT and used 200 sentences as stimuli. There were 100 ‘goat’-biased sentences (10 sentences * 10 VOT values) and 100 ‘coat’-biased sentences (10 sentences * 10 VOT values). The subjects were instructed to press the button labeled ‘yes’ or ‘no’ based on whether the word on the screen matched what they heard. The ‘coat’ response (%) and the response time were used as dependent variables. Results showed the labeling boundary was shifted in favor of the lexical context and peak response time was shifted away from the biased endpoint, suggesting that lexical context influenced categorical perception.

Norris, McQueen, and Cutler (2003) studied how categorical perception is influenced by long-lasting exposure to words with an ambiguous segment. In their study, subjects were divided into two groups and trained to listen to words containing the ambiguous fricative between /f/ and /s/ either in a /f/-favored context (e.g. ‘cliff’ or ‘beef’) or a /s/-favored context (e.g. ‘kiss’ or ‘peace’). After training, subjects changed

their interpretation of an ambiguous fricative in isolation to a lexically biased segment. Such perceptual learning effects continued for at least 12 hours (Eisner & McQueen, 2006; Kraljic & Samuel, 2005). Findings suggest speech perception can be influenced by exposure to a specific linguistic context or environment by showing perception can be adapted and learned.

Comparing word-congruent RTs with nonword-congruent RTs in the boundary and endpoint regions of the continua, Connine and Clifton (1987) found that word-congruent RTs were faster than nonword-congruent RTs though this word RT advantage did not appear in the stimuli at endpoints, and RTs to ambiguous stimuli were slower than RTs to clear ones. The results were interpreted that lexical knowledge could provide an additional information to the representation of the word-congruent phoneme, increasing response times, but the clear phonetic information yielded faster response decisions before lexical information could have a large impact on phoneme processing.

The above studies demonstrate the overt behaviors associated with categorical perception and the effects of context. However, behavioral data to show the lexical influence on categorical perception does not support wholly either an autonomous (feedforward) or interactive (feedback) model. The results can be interpreted as evidence of lexical knowledge feeding back to phoneme processing or lexical knowledge and phoneme processing feeding forward to phoneme identification decision. ERP data may provide information about underlying neurophysiologic processing that may support one model or the other.

Event-related Potentials on Categorical Perception of Speech

Since Sutton, Braren, Zubin, and John (1965) first reported the P300, it has been used to examine cognitive function. The P300 has been thought to reflect sequential information processing, short-term memory, decision making, or attention switching (Squires, Whichens, Squires, & Donchin, 1976). To examine how great a deviant contrast the P300 is sensitive to, that is, how perceptually different stimuli must be to elicit P300, the effect of stimulus frequency, intensity, and duration has been studied (e.g., Cass & Polich, 1997; O'Brien & Stuart, 2001; Sugg & Polich, 1995; Vesco, Bone, Ryan, & Polich, 1993).

Maiste, Wiens, Hunt, Scherg, and Picton (1995) examined how categorical perception is reflected on ERPs. They constructed 2000 stimuli from a nine-step /ba-/da/ continuum. One set of 1000 stimuli was composed of the /ba/ endpoint stimuli at 52% and the other eight stimuli at 48% (6% each). The other set of 1000 stimuli was composed of the /da/ endpoint stimuli at 52% and the other eight stimuli at 48% (6% each). Results showed that N200 was most prominent just beyond the category boundary, thus appeared to be related to the difficulty of the task and the amount of effort involved. The N200-P300 complex was evoked in greater amplitude by more improbable stimuli, suggesting it reflected the phonemic categorization of speech stimuli.

Dehaene-Lambertz (1997) studied the acoustic and phonetic sensitivity of MMN and P300. Control (CO), within-category (WC), and between-category (AC) conditions were designed at the native and non-native (Hindi) phonetic boundary. A sixteen step /ba/-retroflex /Da/ continuum was constructed where Hindi listeners perceive 1-6th sound as /ba/, 7-10th as /da/, and 11-16th as retroflex /Da/ whereas English listeners perceive 1-6th as /ba/ and 7-16th as /da/. Six CV syllables (1st, 5th, 9th for native phonetic contrast; 8th, 12th, 16th for non-native phonetic contrast) were selected as stimuli. Twelve listeners were instructed to listen to the four sequential CV syllables (e.g. /ba/ 5- /ba/ 5 - /ba/ 5 - /ba/ 5 for CO; /ba/ 9 - /ba/ 9 - /ba/ 9 - /ba/ 5 for AC; /ba/ 1- /ba/- 1 /ba/- 1 /ba/ 5 for WC) and press a 'yes' or 'no' button, judging whether the 4th syllable was identical to the preceding three syllables. The results showed for native phonetic contrast, P2 was larger at the inferior frontal and parietal regions and MMN was larger at the superior frontal and temporal regions in AC than in WC with no difference between CO and WC. The P300 was larger at superior frontal, temporal, and parietal regions in AC than in WC and was larger at superior frontal, temporal, and occipital regions in CO than in WC. For non-native phonetic contrast, there was no difference between AC and WC with the larger P300 at the occipital region in CO than in WC. Upon the findings that MMN was larger for the native across-category change, much smaller for the native within-category change, and absent for either type of change in non-native contrast, the author assumed that acoustic parameters cannot explain the results, thus the process of phonemic information stored in sensory memory may be reflected in MMN. Thus, this study showed both MMN and P300 were more sensitive to phonetic contrasts rather than acoustic ones, using across-category and within-category stimuli.

Dalebout and Stack (1999) explored the MMN and P300 in three conditions of most difficult, difficult, and easy discrimination contrasts. From a nine-step /da/ to /ga/ synthetic continuum, the most difficult (7th -9th pair), difficult (subject's individual 2-step pair with the highest discrimination performance), and easy contrast (1st-9th pair) were selected. Twelve subjects' behavioral measure was administered prior to the experiment to select 'difficult' contrast individually. For P300 and MMN, about 500 responses (246 standard, 76 deviant) and 1650 responses (1400 standard, 250 deviant) were collected for each contrast condition, respectively. On the grand average waveform, while P300 was observed clearly on easy contrast, there was no evidence of MMN though the relative negativities followed N1 for all three conditions. The analysis of individual waveform showed that for most difficult, difficult, and easy contrasts, the MMN was elicited in 50%, 25%, and 50% of the subjects whereas the P300 was elicited in 0%, 33%, and 100% of the subjects. They concluded that P300 is elicited by phonetic/phonemic differences, which is consistent with the behavioral response while MMN is elicited by acoustic differences, which shows no relationship with behavioral response. They also pointed out that MMN would have the limitation for the clinical application if it is not identified to all discriminable contrasts.

To investigate whether and how N200/P300 complex are affected by speech versus non-speech contexts, Celsis, Doyon, Boulanouar, Pastor, Demonet & Nespoulous (1999) employed active target detection task. The noise-like phoneme, /f/, was used as the target and presented in a speech and non-speech context. Four CV syllables were

used as distracters (e.g. /fi/, /vi/, /si/, and /zi/) for the speech context and four environmental sounds were used as distracters for the non-speech context (e.g. waterfall, paper crumpling, hooter blast, and glass scratching). For each context of speech and non-speech, 150 stimuli were composed of 30 targets and 120 distracters (30*4) so that the target and four distracters in each context were presented with equiprobability (30/150=20%). Ten subjects were instructed to push a response button whenever they detected the target. With no difference in reaction time in speech and non-speech context, a larger N200 occurred to the non-speech context at the left temporoparietal region and a larger P300 to the speech context at the central region. The larger N200 in the non-speech over speech context did not appear in the right hemisphere. They argued that /f/ remains as a phoneme (speech) rather than a noise in automatic perceptual mechanism, and /f/ was processed as more incongruous stimuli in non-speech context, which caused the larger N200. On the other hand, they postulated that the larger P300 in the speech context could be because four CV distracters were categorized as syllables, thus the target (20%) was easily opposed to syllables (80%) in speech context whereas four environmental sounds were not easily assigned to a category. Stating that the larger N200 in non-speech over speech context was localized laterally and posteriorly compared to the MMN on categorical perception of phoneme, they suggested that phonetic representation of a phoneme was an automatic perceptual process, which may integrate acoustic units into phonological units. The results show the N200 was affected by the speech versus non-speech context.

Tampas, Harkrider, and Hedrick (2005) investigated whether MMN reflects acoustic level only or phonetic level as well in speech processing. They used synthetic consonant-vowel (CV) speech (two within-category stimuli) and nonspeech contrasts (two glides of which frequency ramp matched the formant transitions of the CV). They measured the MMN and P300 with behavioral data. These results showed that the MMN was elicited by the nonspeech only while the P300 was elicited by both the speech and nonspeech with larger amplitude and shorter latency in nonspeech and with smaller amplitude and longer latency in the speech stimuli. They also showed that subjects had better discrimination to nonspeech than speech stimuli. They concluded the acoustic and phonetic information are processed in parallel at the level of MMN up to P300 and behavioral response in that MMN, P300, behavioral response were different in within-category CVs versus nonspeech contrasts.

To examine the lexical effect on auditory speech perception, van Linden, Stekelenburg, Tuomainen, and Vroomen (2007) measured MMN using Dutch words. They designed two oddball conditions using a /t/-word (e.g. *vloot*, meaning 'fleet') and /p/-word (e.g. *hoop*, meaning 'hope'). For the /t/-word oddball paradigm, the standard stimuli was /vlo?/ where the final consonant was a sound halfway between /t/ and /p/, and the deviant stimuli was /vlot/. For the /p/-word oddball paradigm, the standard stimuli was /ho?/ where the final consonant was a sound halfway between /t/ and /p/, and the deviant stimuli was /hot/. When they obtained ERP data from F3, Fz, F4, FC3, FCz, and FC4, the difference waveform between /t/-word and /p/-word showed MMN amplitude was significantly larger in /p/-word. The findings were interpreted that the lexical context

affects speech perception on a pre-lexical level rather than a post-lexical level for phonemic decisions.

Specific components of ERP responses have been found to be sensitive to different aspects of speech and language processing. However, there has not been much ERP research on the lexical effect on categorical perception. In addition to behavioral data, ERP data could be used to reveal the underlying processes occurring when there is a lexical effect on categorical perception and/or where speech and language processing are integrated together. In the current study, it will be determined whether N200 and/or P300 reflect the change in categorical perception with context seen during behavioral labeling tasks. Results will be interpreted relative to autonomous versus interactive models of speech perception.

Research Questions and Hypotheses

The research questions and hypotheses of the present study are as follows:

1. The first research question is whether the ERPs (N200 and P300) reflect the lexical effect on categorical perception of speech. N200 amplitude, P300 amplitude, and N2-P3 peak-to-peak amplitude were hypothesized to be larger in the /pi/ context condition than the /bi/ context condition because the deviant stimuli would be perceived more differently from the standard stimuli in /pi/ context condition. Accordingly, N200 latency and P300 latency were also hypothesized to be shorter in the /pi/ context condition than the /bi/ context condition.
2. The second research question is how the scalp distribution of N200 and P300 appears on the given tasks. N200 amplitude might have more of a centro-parietal topography versus its original scalp distribution due to the lexical effect because previous literature suggests semantic processing occurs in more centro-parietal regions. The P300 amplitude was predicted to have greater amplitude on the centro-parietal region in all conditions as has been reported previously.
3. The third research question is whether behavioral data (labeling, discrimination response accuracy, and discrimination reaction times) reflect the lexical effect on categorical perception of speech. The /bi/ labeling response was predicted to be greater in /bi/ context condition than in /pi/ context condition because the labeling response to an ambiguous sound would be biased toward the real word. The discrimination response was predicted to be more accurate and discrimination reaction times to be faster in /pi/ context condition than /bi/ context condition because the farther the perceptual distance is, the easier the discrimination is, and the more accurate and faster the response is.

4. The fourth research question is whether ERP and behavioral data correlate with each other. N200 amplitude was predicted to correlate with the behavioral data based on the findings that N2b amplitude has a close association with discrimination difficulty (Maiste, Wiens, Hunt, Scherg, & Picton, 1995; Senkowski & Hermann, 2002). P300 amplitude was predicted to correlate with the behavioral data as P300 amplitude has been widely reported to have a direct relationship with behavioral discrimination performance (Dalebout & Stack, 1999; Mccarthy & Donchin, 1981; White, Stuart, & Najem, 2010).

CHAPTER 3. METHODOLOGY

Subjects

Monolingual English speakers who are students at The University of Tennessee served as subjects. Twenty seven subjects (male 9, female 18) who ranged in age from 18 to 36 years (Mean=23.3) participated in the labeling test. Because six of them did not return after the labeling test, twenty one (male 5, female 16) who ranged in age from 18 to 28 (Mean=22.3) completed the EEG and discrimination tests. All subjects were right-handed (Oldfield, 1971) and had no history of audiological, neurological, or psychological/psychiatric disorders nor speech, language, hearing, or learning disorders. Their audiometric thresholds were at or below 15dB HL for octave frequencies between 250 and 8000 Hz (re: ANSI, 1996). This study was approved by University of Tennessee Health Science Center Institutional Review Board.

Stimuli

Stimuli were made for two conditions of /bi/ and /pi/ context words. A male native speaker of English produced /bi/, /stiŋ/ (for /bi/ context word, ‘bee sting’), and /sup/ (for /pi/ context word, ‘pea soup’) in a sound attenuating chamber. The speaker attempted to produce the syllables with the same pitch, loudness, and duration/speed. Recording was done in a quiet room using a high quality microphone (Spher-O-Dyne) held approximately 5 cm from the speaker’s mouth. The microphone output was fed to a preamplifier (Tucker-Davis, Model MA2), then routed to a 16-bit A/D converter (Tucker-Davis, Model DD1). Stimuli were saved as a file sampled at 44.1 kHz (CSRE, version 4.5). The recordings were edited into individual tokens using a waveform manipulation software package and then normalized for uniform intensity so that the loudest part of the waveform was set to a specified amplitude (Adobe Audition, Version 1.5). The durations of /bi/, /stiŋ/, and /sup/ were 355 ms, 468 ms, and 406 ms, respectively.

A seven-step series /bi-/pi/ continuum was made by manipulating voice onset time (VOT) from a naturally produced /bi/. After deleting the consonantal portion of natural /bi/ sound, the burst and aspiration portion of synthetic /pi/ was digitally inserted before voicing portion of natural /bi/ by 8ms to control VOT systematically. Thus, a seven-step series of a /bi-/pi/ continuum was constructed so that the Stimulus 1 (i.e. prototypical /bi/) had VOT of 8 msec and Stimulus 7 (i.e. prototypical /pi/) had VOT of 56 msec (Adobe Audition, Version 1.5). These seven stimuli were inserted in /bi/ and /pi/ context words, respectively (e.g. Stimulus 1-/stiŋ/, Stimulus 1-/sup/), and presented on the two-forced choice labeling test (/bi/ or /pi/) to select the standard and deviant stimuli used for the auditory evoked potential oddball paradigm.

Procedure

Prior to the experimental session, all subjects signed informed consent statements, filled out a case history form, and received a screening test.

Labeling Test

To select the standard and deviant stimuli for an oddball paradigm, a two-forced choice labeling test was administered prior to electrophysiological test. Subjects were seated in a sound-treated booth, facing a computer screen which displayed the letters 'b' and 'p'. They listened to the sounds at a comfortable level (approximately 74~76dB SPL) through headphones and instructed to click the mouse on the letter as they heard. Subjects were given a preliminary test using continuum endpoints (i.e. Stimulus 1 and Stimulus 7) to make sure subjects understand the task and can respond accurately to the stimuli. Subjects participated in the study only after their performance was 100% on the preliminary test. Each stimulus was presented ten times in random order, thus subjects gave 140 responses (7 series * 2 context * 10 times) in total.

Based on the results of the labeling test (**Figure 3-1**), Stimulus 1 and Stimulus 4 were selected as standard and deviant stimuli, respectively, for two context conditions in an oddball paradigm. Stimulus 1 was a prototypical /bi/ sound (i.e. /bi/ from a within-category) and Stimulus 4 was an ambiguous sound between /bi/ and /pi/ (i.e. /?/ from a between-category) which showed the lexical effect the most. Thus, '/bistin/' (standard stimulus) and '/?stin/' (deviant stimulus) were used for /bi/ context condition, and '/bisup/' (standard stimulus) and '/?sup/' (deviant stimulus) were used for /pi/ context condition.

A total of 450 stimuli composed of 369 standard stimuli (81%) and 81 deviant stimuli (19%) were presented in an active oddball paradigm for each context condition. 450 stimuli were divided into three blocks such that each block included 150 stimuli which included 123 standard stimuli (81%) and 27 deviant stimuli (19%). Thus, three blocks were made for /bi/ context condition and another three blocks were made for /pi/ context condition. The stimuli were presented in pseudorandom sequences with at least eight standard stimuli preceding the first deviant and no less than 3 standard stimuli after one deviant. Each stimulus was presented with a 1700 msec ISI to allow ample time for the behavioral response. The time to complete each block was about 4 minutes.

Electrophysiological Test and Discrimination Test

Subjects were seated comfortably in a reclining chair in a dark sound-treated booth, given the response button, and connected to the NeuroScan system. Subjects were told that they would listen to the frequent stimuli of /bi/ in both sting and soup conditions, and instructed to press the response button as soon as they heard a stimulus that sounded

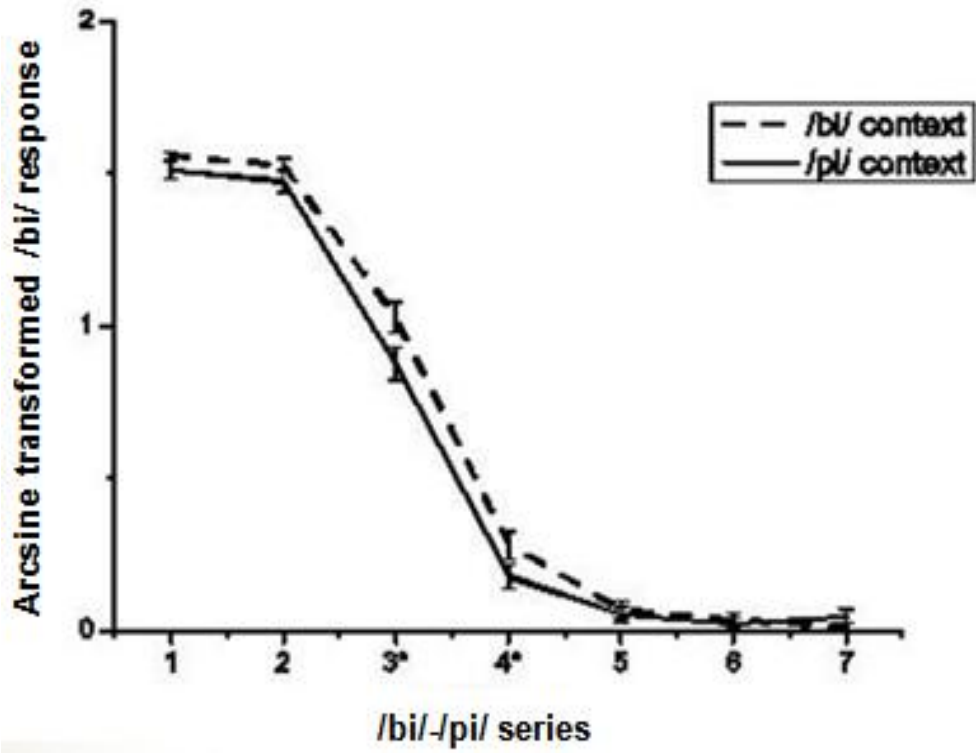


Figure 3-1. /bi/ response in /bi/ and /pi/ contexts by VOT.

Note. Error bars indicate 1 standard error from the mean.

* $p < .05$

different from the frequent stimulus. Subjects were given a 3-minute break every 4 minutes upon the completion of each block. The presentation order of context conditions was counterbalanced across subjects.

The stimuli were presented binaurally, using Etymotic ER-3A insert earphones at 75 dB SPL (± 1 dB), which is consistent with the loudness levels of normal conversational speech. All data were collected with Compumedics NeuroScan hardware and software. For behavioral responses, discrimination response accuracy and discrimination reaction times (from stimulus onset to response onset) were obtained and EEG data associated with an incorrect response were excluded from all analyses off line.

Electroencephalogram (EEG) activity was recorded while subjects were responding behaviorally. EEG responses were acquired from F3, Fz, F4, C3, Cz, C4, P3, Pz, and P4 [re: International 10-20 System (Jasper, 1958)] using a 64-channel NeuroScan system, referenced by the cap electrode between Cz and CPz, and grounded by the cap electrode between Fz and FPz. Vertical and horizontal electrooculograms (EOGs) were recorded by electrodes placed above, below, and on the inner and outer canthi of the left eye. All electrodes were Ag/AgCl, and impedance was maintained at 5 K Ω . The continuous EEG recordings were obtained using a bandpass of 0.01-30 Hz and digitally sampled at 500 Hz for 100 ms before and 1000 ms after the stimulus onset. The remaining trials after excluding EOG greater than 70 μ V were averaged on the basis of their experimental conditions. Subjects who had greater than 58 averaged responses in each context condition were included for further analysis. Grand average waveforms across subjects were derived from averaging all individual waveforms.

Data Analysis

Labeling Test

A two-way repeated measures analysis of variance (ANOVA) was conducted on the arcsine transformed value of the percentage of /bi/ response with the context condition (2 levels: /bi/ context, /pi/ context) and VOT (7 levels: Stimulus 1~7) as the factors. Bonferroni corrections were used for any subsequent univariate testing.

Electrophysiological Test

All post-stimulus (0 – 1000 ms) data points on the deviant waveform were adjusted automatically relative to the average amplitude of the prestimulus baseline (-100 – 0 ms). N1 and P2 in the standard and deviant waveforms were used when necessary to identify N200 and P300. Amplitude and latency of N200 and P300 were measured on the deviant waveform. The most negative peak was scored for N200 between 250-320 ms and for the P300 between 320-500 ms, based on previous research (Celsis, Doyon, Boulanouar, Pastor, Démonet, & Nespoulous, 1999; Maiste, Wiens, Hunt, Scherg, &

Picton, 1995; Tampas, Harkrider, & Hedrick, 2005) and the visual inspection of the group grand average waveform. Five separate repeated-measures ANOVAs were administered, one on N200 latency, one on N200 peak amplitude, one on P300 latency, one on P300 peak amplitude, and one on the N200-P300 complex. The two factors used in the ANOVAs were context condition (2 levels: /bi/ context, /pi/ context) and scalp distribution [frontal (F3, Fz, F4), central (C3, Cz, C4), posterior (P3, Pz, P4)]. Mauchly's test of sphericity was conducted for factors with 3 levels. When significant, Greenhouse-Geisser adjusted degrees of freedom, F- and *p*-values were reported. On the use of multiple ANOVAs, Bonferroni corrections were used such that an alpha level of $p < .001(.05/5)$ was considered and sequential Bonferroni adjustments (Holm, 1979) were used when applicable.

Discrimination Test

Discrimination response accuracy was analyzed on the arcsine transformed value of the percentage of correct responses, and discrimination reaction times (msec) was analyzed on the time between the triggering point of the stimuli and the moment listeners pressed the response button. Two repeated-measures ANOVAs were conducted on the discrimination response accuracy and the discrimination reaction times, respectively, with context condition (2 levels: /bi/ context, /pi/ context) as the factor. Bonferroni corrections were used for any subsequent univariate testing. On the use of multiple ANOVAs, Bonferroni corrections were used such that an alpha level of $p < .025(.05/2)$ was considered and sequential Bonferroni adjustments were used when applicable.

Correlation between Electrophysiological and Behavioral Tests

The correlation between ERP data and behavioral data was analyzed using Pearson product-moment correlations between N200 amplitude, N200 latency, P300 amplitude, P300 latency, N200-P300 peak-to-peak amplitude, discrimination response accuracy, and discrimination reaction times. On the use of multiple bivariate correlations, Bonferroni corrections were used such that an alpha level of $p < .002 (.05/21)$.

CHAPTER 4. RESULTS

Labeling Data

A two-way, repeated measures ANOVA was administered on the arcsine transformed /bi/ response scores as described above. The /bi/ response was significantly different for lexical context ($F_{1,26} = 8.180, p = 0.008, \eta^2p = .239$), VOT ($F_{3,173, 82.490} = 736.782, p < 0.001, \eta^2p = .966$), and the interaction of lexical context and VOT ($F_{3,498, 90.936} = 2.774, p = 0.038, \eta^2p = .096$). Simple effect tests showed that the /bi/ response was greater in /bi/ context than /pi/ context on Stimulus 3 ($F_{1,26} = 5.168, p = 0.032, \eta^2p = .166$) and Stimulus 4 ($F_{1,26} = 6.291, p = 0.019, \eta^2p = .195$). From the labeling results, a prototypical /bi/ sound (Stimulus 1) was selected as a standard, and an ambiguous sound between /bi/ and /pi/ (Stimulus 4) which showed the largest lexical effect was selected as a deviant sound to use in the oddball paradigm. **Figure 3-1** and **Table 4-1** show the results of simple effects for the interaction of lexical context and VOT on the arcsine transformed /bi/ response.

Electrophysiological Data

Figure 4-1 and **Figure 4-2** show the grand average waveforms to the standard and deviant stimuli in /bi/ and /pi/ context conditions. **Figure 4-3** shows the grand average deviant waveforms at Fz magnified in /bi/ context (dashed) and /pi/ context (solid) to indicate each ERP component clearly.

N200

The clear N200 response was elicited by both context conditions at all electrodes. N200 amplitude was significantly different for scalp distribution ($F_{1,218, 25.364} = 40.478, p < 0.001, \eta^2p = 0.669$) but not for lexical context ($F_{1,20} = .011, p = 0.918, \eta^2p = 0.001$) and the interaction of lexical context and scalp distribution ($F_{1,441, 28.811} = .824, p = 0.413, \eta^2p = 0.040$). N200 amplitude was larger on the frontal region than the central region, then the central region than the parietal region. N200 latency was significantly different for scalp distribution ($F_{1,704, 34.075} = 31.015, p < 0.001, \eta^2p = 0.608$) but not for lexical context ($F_{1,20} = 2.264, p = 0.148, \eta^2p = 0.102$) and the interaction of lexical context and scalp distribution ($F_{1,208, 24.170} = 2.116, p = 0.156, \eta^2p = 0.096$). N200 latency was larger on the frontal region than the central region, then the central region than the parietal region. **Table 4-2** shows the results of N200 amplitude and latency by scalp distribution.

P300

A robust P300 response was elicited by both context conditions at all electrodes. P300 amplitude was significantly different for scalp distribution ($F_{1,385, 27.701} = 64.721, p$

Table 4-1. Post-hoc tests for the interaction of lexical context and VOT.

VOT	Contrast pair	<i>p</i>-value
Stimulus 1	/bi/ context vs /pi/ context	.150
Stimulus 2	/bi/ context vs /pi/ context	.205
Stimulus 3	/bi/ context vs /pi/ context	.032*
Stimulus 4	/bi/ context vs /pi/ context	.019*
Stimulus 5	/bi/ context vs /pi/ context	.498
Stimulus 6	/bi/ context vs /pi/ context	.663
Stimulus 7	/bi/ context vs /pi/ context	.083

**p* < .05

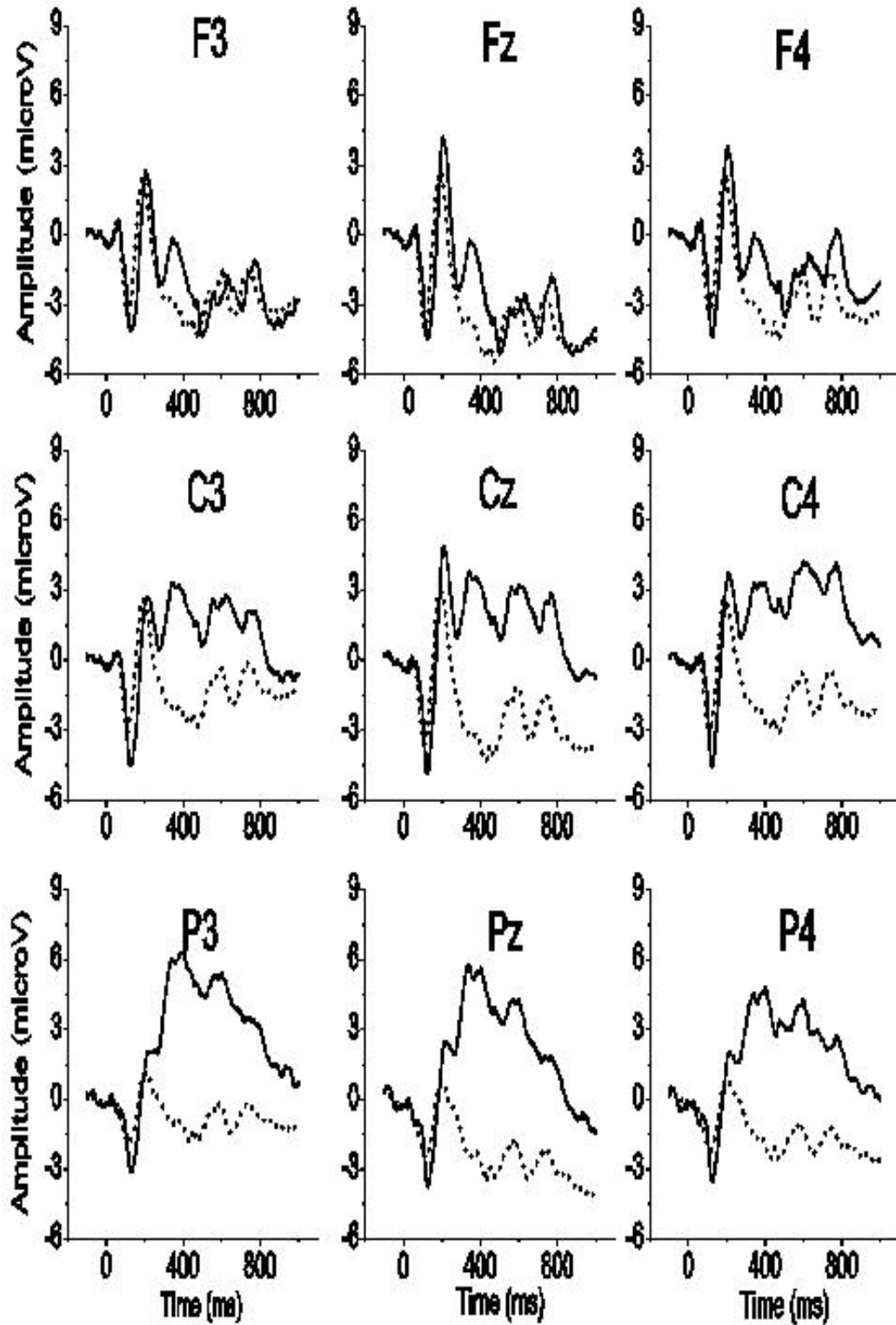


Figure 4-1. Grand average waveforms to the standard stimuli (dashed) and to the deviant stimuli (solid) in /bi/ context.

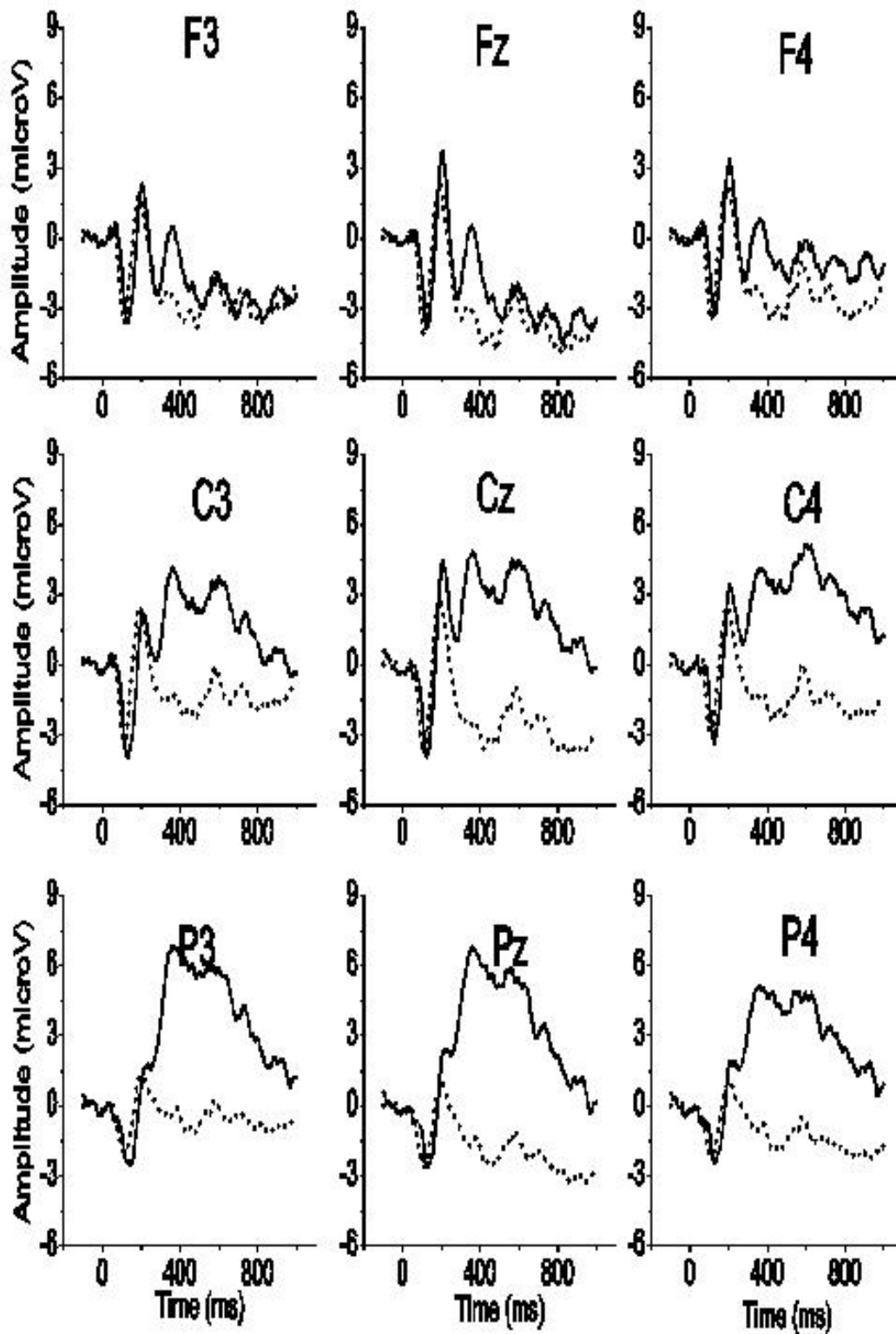


Figure 4-2. Grand average waveforms to the standard stimuli (dashed) and to the deviant stimuli (solid) in /pi/ context.

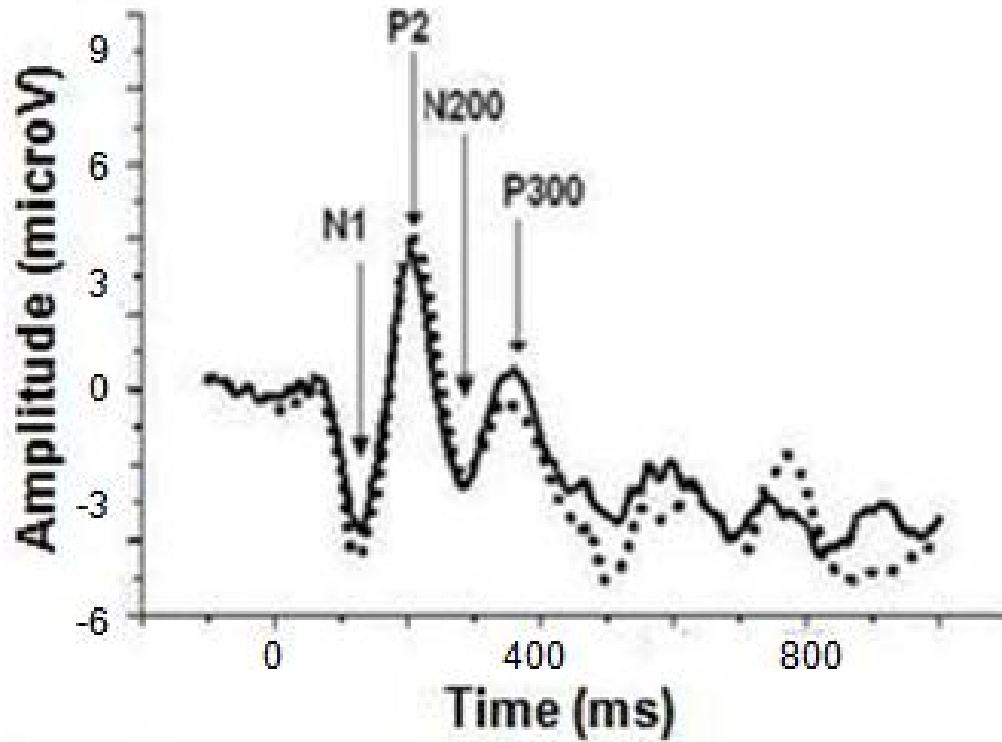


Figure 4-3. Grand average deviant waveforms at Fz magnified in /bi/ context (dashed) and /pi/ context (solid).

Table 4-2. N200 amplitude and latency by scalp distribution.

Contrast pair	<i>p</i> -value
<u>amplitude</u>	
Frontal vs Central	<.001***
Central vs Parietal	.007**
Frontal vs Parietal	<.001**
<u>latency</u>	
Frontal vs Central	0.013*
Central vs Parietal	<.001***
Frontal vs Parietal	<.001***

****p* <.001, ***p* <.01, **p* <.05

< 0.001, $\eta^2p=0.764$) but not for the lexical context ($F_{1,20} = 1.103, p = 0.306, \eta^2p = 0.052$), or the interaction of lexical context and scalp distribution ($F_{1.336, 26.722} = 0.361, p = 0.615, \eta^2p = 0.018$). The P300 amplitude was larger on the posterior region than the central region, and greater on the central region than the frontal region. P300 latency was not significantly different for the lexical context ($F_{1,20} = 0.849, p = 0.368, \eta^2p = 0.041$), scalp distribution ($F_{1.488, 29.762} = 2.50, p < 0.112, \eta^2p = 0.111$) and the interaction of lexical context and scalp distribution ($F_{2,40} = 0.693, p = 0.506, \eta^2p = 0.033$). **Table 4-3** shows the result of P300 amplitude by scalp distribution.

N200-P300 Complex

A N200-P300 complex was elicited by both context conditions at all electrodes. N2-P3 peak-to-peak amplitude was significantly different for scalp distribution ($F_{1.361, 27.701} = 13.323, p < 0.001, \eta^2p = 0.400$) but not for the lexical context ($F_{1,20} = 2.484, p = 0.131, \eta^2p = 0.110$), or the interaction of lexical context and scalp distribution ($F_{1.281, 25.615} = 1.354, p = 0.265, \eta^2p = 0.063$). The N200-P300 peak-to-peak amplitude was larger on the posterior region than the central region, and greater on the central region than the frontal region. **Table 4-4** shows the result of N200-P300 complex by scalp distribution.

Discrimination Data

Two one-way, repeated measure ANOVAs were administered, one on the arcsine transformed discrimination response accuracy score and one on discrimination reaction times, with lexical context (2 levels: /bi/ context, /pi/ context) as the factor. Discrimination response accuracy was not significantly different for lexical context ($F_{1,20} = 1.701, p = 0.207, \eta^2p = 0.078$). Discrimination reaction times were significantly different ($F_{1,20} = 4.946, p = 0.038, \eta^2p = 0.198$) with sequential Bonferonni adjustments such that they were faster in the /pi/ versus /bi/ context. **Table 4-5** shows the mean and standard deviation of response accuracy and reaction times by individual and group. The behavioral responses on the lexical context are shown in **Figure 4-4**.

Correlation between EEG and Behavioral Data

There was no relationship between behavioral and ERP response. P300 amplitude was positively correlated with N200 amplitude and N200-P300 peak-to-peak amplitude while N200-P300 peak-to-peak amplitude was negatively correlated with N200 latency. **Table 4-6** shows the results of the bivariate correlation with Pearson correlation coefficient.

Table 4-3. P300 amplitude by scalp distribution.

Contrast pair	<i>p</i>-value
Frontal vs Central	<.001*
Central vs Parietal	<.001*
Frontal vs Parietal	<.001*

**p* <.001

Table 4-4. N2-P3 complex by scalp distribution.

Contrast pair	<i>p</i>-value
Frontal vs Central	<.034*
Central vs Parietal	<.004*
Frontal vs Parietal	<.002*

**p* <.001

Table 4-5. Discrimination response accuracy and reaction times by subject.

Subject	<u>Response accuracy (%)</u>		<u>Reaction times (ms)</u>	
	/bi/ context	/pi/ context	/bi/ context	/pi/ context
S1	81	81	553.8	467.5
S2	64	65	378.5	423.9
S3	81	81	306.8	305.3
S4	80	77	817.7	689.2
S5	79	72	528.7	526.5
S6	77	79	514.8	446.6
S7	81	80	713.6	685.5
S8	76	76	378.4	393.1
S9	81	81	701.9	678.0
S10	80	80	505.4	478.1
S11	64	74	659.3	594.7
S12	60	56	588.1	574.6
S13	68	75	515.6	518.8
S14	79	81	574.3	501.2
S15	78	79	646.7	614.8
S16	57	73	842.9	660.6
S17	80	81	410.1	357.0
S18	78	81	459.0	421.3
S19	79	80	713.9	682.3
S20	69	62	473.6	586.1
S21	79	80	438.7	482.0
Mean	74.8	75.9	558.2	528.0
SD	7.7	7.0	145.8	115.6

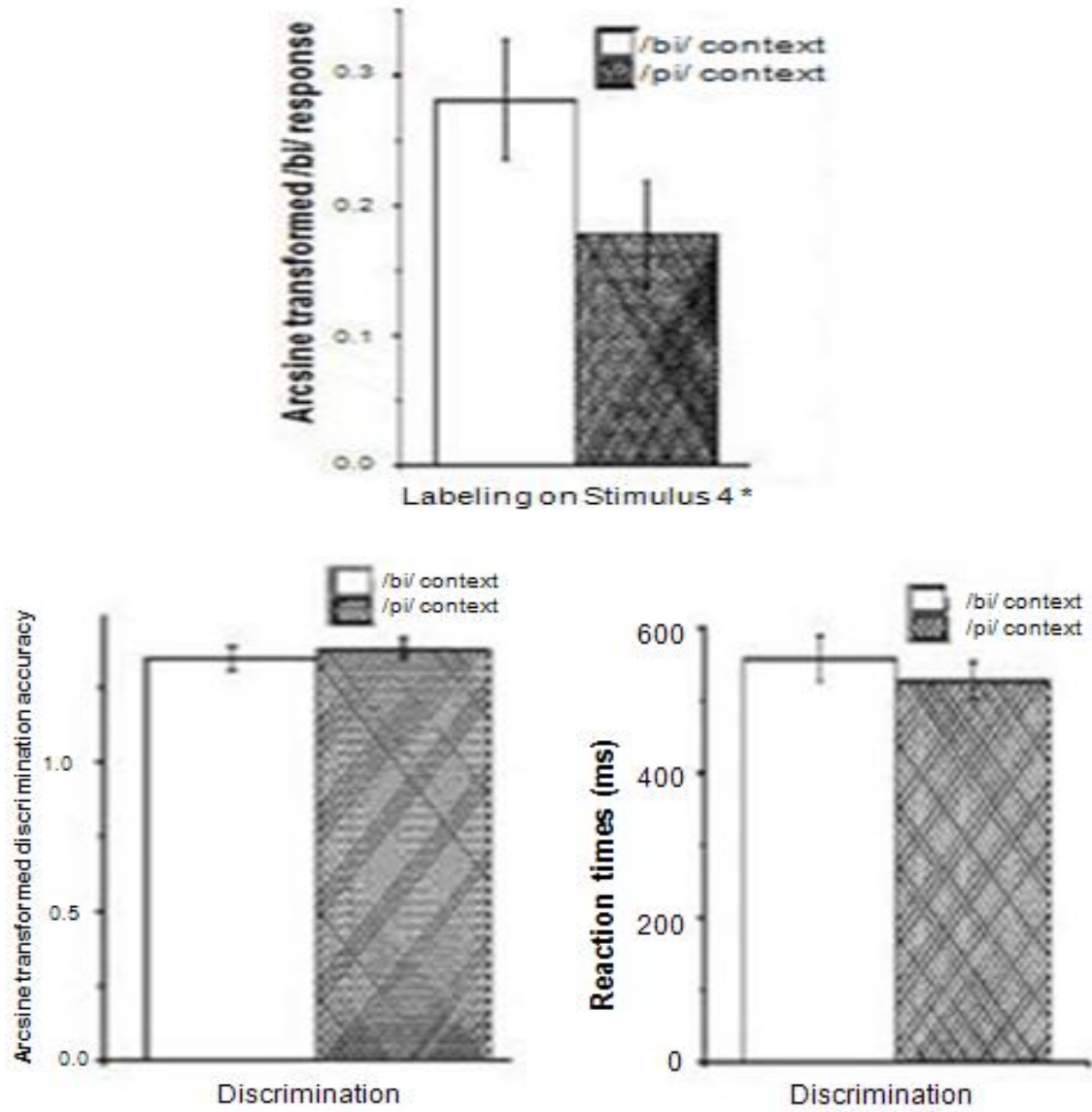


Figure 4-4. Lexical effect on behavioral responses.
Note. Error bars indicate 1 standard error from the mean.
 * $p < .05$

Table 4-6. Bivariate correlation in /bi/ and /pi/ context conditions.

Correlation	N2 amp^a	N2 lat^b	P3 amp^c	P3 lat^d	N2-P3^e	Accuracy^f	RT^g
N2 amp	1.0	.125	.518*	-.056	-.356	.164	-.007
N2 lat		1.0	-.447	.253	-.604*	-.247	.142
P3 amp			1.0	-.220	.615*	.411	-.426
P3 lat				1.0	-.189	.032	.155
N2-P3					1.0	.298	-.458
Accuracy						1.0	-.168
RT							1.0

* $p < .001$

^a N200 amplitude

^b N200 latency

^c P300 amplitude

^d P300 latency

^e N200-P300 peak-to-peak

^f Discrimination response accuracy

^g Discrimination reaction times

CHAPTER 5. DISCUSSION

Interpretation of the Current Results

The current study explores whether ERP components (N200 and P300) reflect the influence of lexical context on categorical perception of speech.

The first question of the present study was whether N200 and P300 reflect the lexical effect on categorical perception of speech. N200 and P300 are the ERP components associated with a pre-lexical level of processing (acoustic or phonetic) though there are some findings to show they also reflect the lexical level. Different from the hypotheses, current results show N200 amplitude, P300 amplitude, N200 latency, P300 latency, and N200-P300 peak-to-peak amplitude do not demonstrate a change in categorical perception in the presence of lexical context. van Lindern, Stekelenburg, Tuomainen, and Vroomen (2007) found that MMN shows a lexical effect on categorical perception. They suggested the results were the evidence of lexical effects occurring on the pre-lexical rather than the post-lexical level.

If the lexical context is reflected on the MMN, it is more likely to be reflected on the N200 and P300. Such inconsistent results might be possibly due to the small effect size of the lexical context in the present study. However, when 15 subjects who showed clear lexical effects on the labeling test were analyzed separately, their N200 and P300 data were not also significantly different. In addition, while both are obtained using oddball paradigms, MMN is pre-attentive and pre-conscious while attention is required to evoke the N200 and P300. Thus, there are distinct differences between MMN and N200/P300 which could account seemingly conflicting findings. It is still elusive whether the lexical effect on speech perception occurs on the pre-lexical level or post-lexical level. Future research would need to understand the neurophysiologic mechanism.

Though N200 did not reflect a shift in the perception of deviant stimuli based on context, a robust N200 was elicited by all deviant stimuli as in other previous studies. Previous findings suggest N200 is related to cognitive processes like decision making related to the behavioral discrimination task (Folstein & Van Petten, 2008; Ritter et al., 1979). Given that between-category stimuli were used as deviant stimuli, the results also suggest N200 is sensitive to acoustic or sub-phonetic processing. The findings that N200 amplitude and latency were not different for the lexical context seem to be in line with the previous studies about N200 or PMN (phonological mapping negativity). PMN is N200 but named to represent the nature of a phonological process, suggesting it reflects an autonomous phonological stage rather than an early lexical selection stage of word processing. Much research has reported that N200/PMN is elicited in both phonological match and mismatch conditions in a priming paradigm, reflecting prelexical processing of phonological activation (Connolly et al, 2001; D'Arcy, Connolly, Service, Hawco, & Houlihan, 2004; Lee, Harkrider, & Hedrick, 2012). Some studies suggested N200 is insensitive to acoustic processing (Ritter, Paavilainen, Lavikainen, Reinikainen, Alho, Sams, & Näätänen, 1992; Schröger, 1996), unlike the MMN. It might be dependent on

the extent of perceptual difference of acoustic change. From the findings of current and previous studies, N200 is thought to be sensitive not only to phonetic but also to acoustic or sub-phonetic change of speech sounds, but is not involved with lexical context.

Similar to N200, P300 did not reflect the lexical effect on categorical perception of speech stimuli in context, though P300 also was clearly elicited by the deviant stimuli. It has been known that the P300 amplitude is very sensitive to stimulus probability and task (Folstein & Van Petten, 2008). Given that between-category stimuli were used as deviant stimuli, the results suggest P300 is sensitive to acoustic or sub-phonetic processing as well. Because P300 has shown good correlation with behavioral performance (Dalebout & Stack, 1999; White, Stuart, & Najem, 2010), it may have been so robust because the deviant stimuli were easily discriminated from the standard stimuli, behaviorally. For example, using auditory stimuli, White, Stuart, and Najem (2010) reported P300 was evoked by behaviorally perceptible contrasts whereas no P300 was evoked by behaviorally imperceptible contrasts. Though there has been some controversy about whether P300 reflects acoustic or phonetic processing, current data may suggest P300 is not only to phonetic but also to acoustic or sub-phonetic change of speech sound, but not be associated with lexical context.

The second question of the present study was how the scalp distribution of N200 and P300 appears during the given tasks. Differently from hypothesis, N200 showed the original scalp distribution with larger amplitude on the frontal, then central, then parietal electrodes. It is in line with previous findings from studies that did not include lexical factors, such that N200 is dominant over the frontal or frontocentral region (Folster & Van Petten, 2008; Patel & Azzam, 2005). Also, N200 latency was longer on the frontal, then central, then parietal electrodes, suggesting the larger N200 amplitude, the longer N200 latency. Because introducing linguistic context to a speech task did not alter scalp distribution of N200, these findings further support the suggestion that N200 does not reflect the lexical effect on categorical perception of speech stimuli. In addition, as hypothesized, P300 amplitude was greater on the parietal, then the central, then the frontal region. P300 latency did not change depending on the scalp distribution. Findings are consistent with previous data such that the P300 is more dominant over centroparietal regions (Folstein & van Petten, 2008; Martin, Tremblay, & Korczak, 2008; Picton, 1992) and the amplitude of P300 is a more sensitive indicator than latency. Results suggest the P300 response to speech under lexical effect is the same as P300 response to speech stimuli without lexical effect. Because the scalp of P300 originally is centroparietal region where semantic processing also occurs as suggested in N400 (Holcomb & McPherson, 1994; Kutas & Hillyard, 1980), the unaltered scalp distribution of P300 could not be interpreted to support that P300 is not associated with lexical effect.

The third question of the present study was if labeling, discrimination response accuracy, and discrimination reaction times reflect lexical effect on categorization of speech stimuli. The labeling and discrimination reaction times, as hypothesized, reflected the lexical effect whereas discrimination response accuracy did not, differently from expectation. The /bi/ labeling response was greater in the /bi/ context condition than /pi/ context condition. It is consistent with the previous studies' findings reporting Ganong

effect (Borsky, Tuller, & Shapiro, 1998; Ganong, 1980; Pitt, 1995). Also, the discrimination reaction times were faster in /pi/ context condition than /bi/ context condition. This may be because the perceptual distance between the standard and deviant stimuli were farther in /pi/ context condition than in /bi/ context condition, thus, the discrimination task was easier in /pi/ context condition than in /bi/ context condition. This finding shows the faster reaction times are associated with less processing demand, but also shows the lexical effect as well. The discrimination response accuracy, however, was not sensitive to the lexical effect. It may be because it was a very easy task to discriminate the deviant stimuli from the standard stimuli, so it caused the ceiling effect on the subjects' response. Thus, discrimination response accuracy was not sensitive to evaluate the subjects' behavioral performance.

The fourth question of the present study was if ERP and behavioral data correlate with each other. There was no correlation between ERP components and behavioral response. Though P300 amplitude has been thought to have great association with behavioral response, it seems because P300 amplitude and N200-P300 peak-to-peak amplitude did not vary in wide range depending on the lexical effect, which caused no correlation with behavioral responses (e.g. discrimination response accuracy or discrimination reaction times).

The above findings can be interpreted in the models of speech perception. N200 and P300 did not reflect the lexical context effect which behavioral data (in labeling and discrimination reaction times) showed. Traditionally, Ganong effect on categorical perception has been interpreted as a feedback process from the higher lexical level to lower pre-lexical levels in an interactive or top-down model (e.g. TRACE; McClelland & Elman, 1986). However, lexical effects on categorical perception can be also interpreted as feedforward processing from the pre-lexical and lexical level to the decision making level in autonomous or bottom up models without feedback process (e.g. MERGE; Norris, McQueen, & Cutler, 2001). The current data show N200 and P300 did not reflect the lexical context effect on categorical perception of speech and have the same scalp distribution as in no lexical condition. Thus, the results are more in line with autonomous model, suggesting that the lexical effect does not affect the pre-lexical level.

Limitations and Suggestion for Future Research

Stimulus 4 was selected as deviant stimuli because the most ambiguous sound between categories could bolster the lexical effect on categorical perception. As discussed above, however, it was a very easy task to discriminate the deviant stimuli (Stimulus 4) from the standard stimuli (Stimulus 1), which caused discrimination response accuracy not to reflect lexical effect. Further, not only behaviorally but also electrophysiologically, the task might have been too easy to reveal the lexical effect. The task difficulty should be carefully considered in measuring endogenous responses.

The stimuli for two context conditions were divided into six blocks (three in /bi/ context condition, three in /pi/ context condition). Though the order of presentation was

counterbalanced to avoid the order effect, one condition and the other condition are tested separately, thus the extent of subjects' motivation could not be controlled block by block as much as two conditions are included in one block. It would not occur if two conditions were randomized in one block. It was inevitable to include two context conditions with two different standard stimuli, but order effects and subjects' attention/motivation need to be well controlled when more than one condition is considered.

Though the sample size of 21 was decided based on a power analysis with the effect size of about .2, similar studies have not reported their effect size so a larger sample size may be necessary to show the lexical effect on N200 or P300 in the present study.

The deviant stimulus (Stimulus 4) was selected based on the labeling results, reflecting Ganong effect. However, it was discrimination response that was required in an oddball paradigm. Thus, it may show some direct relationship between behavioral and electrophysiological effect if labeling task is used in an oddball paradigm.

EEG has a non-invasive procedure with high temporal resolution, and no exposure to high-intensity magnetic fields and radioligands. Specially, ERPs seem to be an efficient tool to examine neurophysiologic processes underlying speech and language processing. Though N200 and P300 have been thought to be associated with higher cognitive processing, it is not clear whether they reflect lexical/semantic processes. Speech perception is very complicated involving with both sensory (physical) and cognitive (linguistic) factors (van Rooij & Plomp, 1991). ERP research on speech perception associated with lexical component can provide the integrative information about speech and language processing for research or/and clinical purpose. The better neurophysiologic understanding of speech and language processing could contribute to model/design speech recognition or deciding signal processing for amplification devices. Gatehouse et al. (2006) reported that people with poor cognitive function showed better speech performance (i.e. speech recognition in noise) with slow-acting compression while people with good cognitive function showed better speech performance with fast-acting compression. Thus, in addition to general cognition, more specified ERP data could be used as a good indicator to decide the most fit signal processing. In addition, ERP data on speech-language processing could apply for clinical diagnosis, treatment, and prognosis as well. For instance, ERP data could be helpful to develop speech-language perception battery to monitor the fitting of hearing aids or the pre-and post-cochlear implant operation results, and to decide what is the most fit signal processing or how well the speech-language treatment is going.

Conclusion

The present study shows that N200 and P300 do not reflect the previously documented behavioral effect of a strong lexical influence on categorical perception. Findings suggest that the lexical context does not affect the pre-lexical level, alternatively

supporting the autonomous model of speech perception where lexical context occurs on the post-lexical level.

LIST OF REFERENCES

- Aaltonen, O., Paavilainen, P., Sams, M., & Näätänen, R. (1992). Event-related brain potentials and discrimination of steady-state vowels within and between phoneme categories: A preliminary study. *Logopedics Phoniatrics Vocology, 17*, 107-112.
- Abramson, A. S. (1979). The noncategorical perception of tone categories in Thai. In B. Lindblom and S. Öhman B. (Eds.), *Frontiers of speech communication research* (pp. 127-134). London: Academic Press.
- Bennington, J. Y., & Polich, J. (1999). Comparison of P300 from passive and active tasks for auditory and visual stimuli. *International Journal of Psychophysiology, 34*, 171-177.
- Borsky, S., Tuller, B., & Shapiro, L. P. (1998). "How to milk a coat": the effects of semantic and acoustic information on phoneme categorization. *Journal of Acoustical Society of America, 103*, 2670-2676.
- Borstein, M. H. & Korda, N. O. (1984) Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research, 46*, 207-222.
- Cass, M. & Polich, J. (1997). P300 from a single stimulus paradigm: auditory intensity and tone frequency effects. *Biological Psychiatry, 46*, 51-64.
- Celsis, P., Doyon, B., Boulanouar, K., Pastor, J., Démonet, J. F., & Nespoulous, J. L. (1999). ERP correlates of phoneme perception in speech and sound contexts. *Neuroreport, 14*, 1523-1527.
- Conchin, C. C. & Coles, M. G. H. (1988) "Precommentary: Is the P300 component a manifestation of context updating?" *Behavioral and Brain Science, 11*, 355-425.
- Connine C. M. & Clifton, C. (1987). Interactive use of lexical information in speech Perception. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 291-299.
- Connolly, J. F., Service, E., D'Arcy, R. C. N., Kujala, A., & Alho, K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *NeuroReport, 12*, 237-243.
- Cummings, A., Ceponiene, R., Koyama, A. L, Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115*, 92-107.

- Cutler A., Mehler J., Norris D., & Segui J. (1987). Phoneme identification and the lexicon. *Cognitive Psychology*, *19*, 141–177.
- D’Arcy, R. C. N., Connolly, J. F., Service, E., Hawco, C. S., & Houlihan, M. E. (2004). Separating phonological and semantic processing in auditory sentence processing: A high-resolution event-related brain potential study. *Human Brain Mapping*, *22*, 40-51.
- Dalebout, S. D., & Stack, J. W. (1999). Mismatch negativity to acoustic differences not differentiated behaviorally. *Journal of the American Academy of Audiology*, *10*, 388-399.
- Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport*, *8*, 919-924.
- Donchin, E. (1981). Surprise! . . . Surprise? *Psychophysiology*, *18*, 493–515.
- Donchin, E. & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 357-374.
- Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time. *Journal of the Acoustical Society of America*, *119*, 1950-1953.
- Etcoff, N. L. & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*, 227-240.
- Evans, J. L., Selinger, C., & Pollak, S. D. (2011). P300 as a measure of processing capacity in auditory and visual domains in specific language impairment. *Brain Research*, *1389*, 93-102.
- Folstein, J.R. & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*, 152-170.
- Frenck-Mestre, C., Meunier, C., Espesser, R., Holcomb, P., & Daffner, K. (2005). Perceiving nonnative vowels: The effect of context on perception as evidenced by event-related brain potentials. *Journal of Speech, Language, and Hearing Research*, *48*, 496-1510
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 110-125.

- Gatehouse, S., Naylor, G., Elberling, C. (2006). Linear and non-linear hearing aid fittings-2. Patterns of candidature. *International Journal of Audiology*, 45, 153-171.
- Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: A critical overview. In S. Harnad (Ed.), *Categorical Perception: The Groundwork of Cognition* (pp. 1-25). New York: Cambridge University Press.
- Hillyard, S. A. & Picton, T. (1987). Electrophysiology of cognition. In F. Plum (Ed.), *Handbook of physiology: Section I. The nervous system: Vol. 5. Higher functions of the brain, part 2* (pp. 519-584). Bethesda, MD: Waverly Press.
- Holcomb, P.J., McPherson, W.B., 1994. Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition*, 24, 259–276.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.
- Horev, N., Most, T., & Pratt, H. (2007). Categorical Perception of Speech (VOT) and Analogous NonSpeech (FOT) signals: Behavioral and electrophysiological Correlates, *Ear and Hearing*, 28, 111-128.
- Kraljic, T. & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, 51, 141-178.
- Kutas, M., DeLong, K. A., & Smith, N. J. (2011). A look around at what lies ahead: Prediction and predictability in language processing. In M. Bar (Ed.), *Predictions in the Brain: Using Our Past to Generate a Future* (pp. 190-207). Oxford University Press.
- Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, 4, 463–470.
- Kutas, M. & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207, 203-208.
- Lee, J. Y. (2012). Auditory P300 responses evoked by perceptually different between-category CV stimuli. *Korean Journal of Communication Disorders*, 17, 120-128.
- Lee, J. Y., Harkrider, A. W., & Hedrick, M. S. (2012). Electrophysiological and behavioral measures of phonological processing of auditory nonsense V-CV-VCV. *Neuropsychologia*, 50, 666-673.
- Liberman, A. M., Harris, K. S., Hoffman, H. S., and Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358-368.

- Maiste, A. C., Wiens, A. S., Hunt, M. J., Scherg, M., & Picton, T. W. (1995). Event-related potentials and the categorical perception of speech sounds. *Ear and Hearing, 16*, 68-90.
- Martin, B. A., Tremblay, K., & Korczak, P. (2008). Speech evoked potentials: from the laboratory to the clinic. *Ear and Hearing, 29*, 285-313.
- McCarthy, G. & Donchin, E (1981). A metric for thought: a comparison of P300 latency and reaction time. *Science, 211*, 77-80.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*, 1-86.
- McClelland, J. L., Mirman, D. & Holt, L. L. (2006) Are there interactive processes in speech perception ? *TRENDS in Cognitive Sciences, 10*, 363-369.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Valnio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language specific phoneme representations revealed by electric and magnetic brain responses. *Nature, 385*, 432-434.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology, 118*, 2544 –2590.
- Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*, 299–370.
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*, 204–238.
- O'Brien, P. J. & Stuart, A. (2001). The effect of auditory stimulus duration on the P300 response. *Journal of Speech-Language Pathology and Audiology, 25*, 19-23.
- Patel, S. H. & Azzam, P. N. (2005). Characterization of N200 and P3000: selected studies of the event-related potential. *International Journal of Medical Sciences, 2*, 147-154.
- Peltola, M. S., Kujala, T., Tuomainen, J., Ek, M., Aaltonen, O., & Näätänen, R. (2003). Native and foreign vowel discrimination as indexed by the mismatch negativity (MMN) response. *Neuroscience Letters, 352*, 25-28.
- Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology, 9*, 456–479.

- Pitt, M. A. (1995). The locus of the lexical shift in phoneme labeling. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1037-1052.
- Pitt, M. A., & Samuel, A. G. (1993). An empirical and metaanalytic evaluation of the phoneme identification task. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 699-725.
- Polich, J. (2007). Updating P300: An integrative theory of P300a and P300b. *Clinical Neurophysiology*, *118*, 2128-2148.
- Potts, G. F., Dien, J., Hartry-Speiser, A. L., McDougal, L. M., & Tucker, D. M. (1998). Dense sensor array topography of the event-related potential to task-relevant auditory stimuli. *Electroencephalography and Clinical Neurophysiology*, *106*, 444-456.
- Pritchard, W., S., Shappell, S.A., & Brandt, M.E. (1991). Psychophysiology of N200/N400: A review and classification scheme. In J.R.Jennings & P.K.Ackless (Eds.). *Advances in psychophysiology: A research annual (Vol 4*, pp. 43-106). London: Jessica Kingsley.
- Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M. and Näätänen, R. (1992). Event-related potentials to repetition and change of auditory stimuli. *Electroencephalography and Clinical Neurophysiology*, *83*, 306-321.
- Ritter, W., Simson, R., Vaughan Jr., H.G., Friedman, D. (1979). A brain event related to the masking of a sensory discrimination. *Science*, *203*, 1358-1361.
- Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research*, *248*, 48-59.
- Sams, M., Aulanko, R. Aaltonen, O., & Näätänen, R. (1990). Event-related potentials to infrequent changes in synthesized phonetic stimuli. *Journal of Cognitive Neuroscience*, *2*, 344-357.
- Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology*, *37*, 473-484.
- Schröger, E. (1996). The influence of stimulus intensity and inter-stimulus interval on the detection of pitch and loudness changes. *Electroencephalography and Clinical Neurophysiology*, *100*, 517-526.
- Sharma, A. & Dorman, F. (1999). Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *Journal of the Acoustical Society of America*, *106*, 1078-1083.

- Sharma, A., Kraus, N., McGee, T., Carrell, T., & Nicol, T. (1993). Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potential. *Electroencephalography and Clinical Neurophysiology*, 88, 64-71.
- Squires, K. C., Wichens, C., Squires, N. K., & Dochin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science*, 193, 1142-1146.
- Sugg, M. J. & Polich, J. (1995). P300 from auditory stimuli: intensity and frequency effects. *Biological Psychology*, 41, 255-269.
- Sutton, S., Braren, M., Zubin, J., & John, E.R. (1965). Evoked-Potential Correlates of Stimulus Uncertainty. *Science*, 150, 1187-1188.
- Tampas, J.W., Harkrider, A.W., Hedrick, M.S. (2005). Neurophysiological indices of speech and nonspeech stimulus processing. *Journal of Speech, Language, and Hearing Research*, 48, 1147-1164.
- van Linden, S., Stekelenburg, J. J., Tuomainen, J., & Vroomen, J. (2007). Lexical effect on auditory speech perception: An electrophysiological study. *Neuroscience Letters*, 420, 49-52.
- van Rooij, J. C. G. M. & Plomp, R. (1991). Auditive and cognitive factors in speech perception by elderly listeners. *Acta Oto-laryngologica (stockholm)*, 476, 177-191.
- Vesco, K. K., Bone, R. C., Ryan, J. C., & Polich, J. (1993). P300 in young and elderly subjects: auditory frequency and intensity effects. *Electroencephalography and Clinical Neurophysiology*, 88, 302-308.
- Wang, J., Miyazato, H., Randall, M., Hokama, H., Hiramatsu, K., & Ogura, C. (2003). The N200 abnormalities of auditory event-related potentials in patients with panic disorder. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 27, 1013-1021.
- White, L., Stuart, A., & Najem, F. (2010). Mismatch negativity and P300 to behaviorally perceptible and imperceptible temporal contrasts. *Perceptual and Motor Skills*, 110, 1105-1118.
- Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csepe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A., & Näätänen, R. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, 7, 357-369.

Young, A. W., Rowland, D., Calder, A. J., Ectoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*, 271-313.

VITA

Ji Young Lee was born in Seoul, Korea in 1973. She received her bachelor's degree in Korean Language and Literature from the Ewha Woman's University and her Master of Science in Interdisciplinary Program of Communication Disorders from the Ewha Woman's University Graduate School. With acceptance of this dissertation, she will receive a Ph.D. degree in Speech and Hearing Science from the University of Tennessee Health Science Center.