EARLY AND LATE STAGES OF NEURAL SPEECH PROCESSING IN NATIVE-ENGLISH AND NATIVE-POLISH LISTENERS: A BEHAVIORAL AND ERP STUDY

Submitted by
Monica Wagner

A dissertation submitted to the Graduate Faculty in Speech-Language-Hearing Sciences
in partial fulfillment of the requirements for the degree of Doctor of Philosophy,
The City University of New York

2011

This manuscript has been read and accepted for the

Graduate Faculty in Speech-Language-Hearing Sciences in satisfaction of the dissertation

requirement for the degree of Doctor of Philosophy

Date	Valerie L. Shafer, PhD Chair of Examining Committee
Date	Klara Marton, PhD Executive Officer

Supervisory Committee:

Brett Martin, PhD

Mitchell Steinschneider, MD, PhD

THE CITY UNIVERSITY OF NEW YORK

Abstract

Early and late stages of neural speech processing in native-English and native-Polish listeners: A behavioral and ERP study

by Monica Wagner

Advisor: Dr. Valerie L. Shafer

The effect of exposure to the contextual features of the consonant cluster /pt/ on speech perception was investigated in native-English and native-Polish listeners using behavioral and ERP methodology. English and Polish listeners experience the consonant cluster /pt/ in their languages, but only the Polish group experiences the cluster in the context examined in the current experiment (i.e., word onset). Acoustic features of phonemes change with context and, therefore, only the Polish listeners are exposed to the acoustic features of the word onset /pt/ cluster. The /st/ cluster occurs in both English and Polish in word onset and, therefore, the /st/ cluster was used as an experimental control. Also, because the /st/ and /pt/ clusters have different acoustic characteristics, the influence of these characteristics on speech perception, irrespective of native-language exposure was examined.

Two and three-syllable nonsense words beginning with /pt/, pət/, /st/ and /sət/ were presented within same and different word pairs. A syllable identification task performed in response to the second word in the pair revealed that Polish listeners were able to distinguish /pt/ and /pət/ nonsense words but English listeners were not, suggesting that contextual features of phonemes are intrinsic to perceptual speech processing. ERP responses to the second word in the pair revealed native-language speech perception to be reflected late in latency within the LPC and P400. Sensory-obligatory responses from fronto-temporal sites (P1N1P2 complex) in response to the first word in the pairs revealed auditory signature waveforms for the 2 and 3-syllable *pt* word forms that were the highly similar for the English and Polish listeners. In contrast, English and Polish groups showed different sensory-obligatory responses from lateral-temporal sites (T-Complex) to the 2-syllable *pt* word forms reflecting native-language speech perception and these differences began early in cortical processing at 40 ms. Together, these findings suggest that both acoustic and linguistic distinctions are reflected at both early and late stages of cortical speech processing but from different brain sources. Also, neurophysiological differences in response to the acoustic characteristics of the /st/ and /pt/ stimuli, irrespective of native language, were evident.

Acknowledgments

Valerie Shafer, Mitchell Steinschneider and Brett Martin
This process has been an incredible privilege

Valerie, thank you for sharing your creative intelligence and gift for writing
Your kindness has made doctoral education fun
There is not a better lab anywhere

Thank you Mitch for all I have learned in the area of neurophysiology of speech
Please allow me to continue

Thank you Brett for introducing me to auditory word signatures of the ACC,
an area of research so important for the communicatively impaired populations

Ayda Tekok-Kilic, I appreciate, very much, your time and advice

---

To the linguists
Katarzyna Dziubalska-Kolaczyk, Malgosia Pyrzanowski, and Zosia.Stankiewicz
Your knowledge concerning
the phonotactic rules of the Polish language has been extremely helpful

To my friends and colleagues at The Graduate Center
Only with your help, support and camaraderie has this been possible
A special thanks to Hia Datta, Yan H.Yu, Gary Chant and Loretta Walker

Dedication

I thank my husband, Peter Wagner, for a lifetime of support
And my father, Bernard Palmieri, for his encouragement

To my children, John, Julia and Meredith
You give my life meaning

To my family, Mom, Anita and Angelo
I love you

Table of Contents

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

A major issue of our time in the neuroscience of speech and language concerns the manner in which the acoustic signals of speech are integrated and translated into cortical processes contributing to speech perception. The study of speech perception of native-language groups permits a natural experiment because speech exposure of native-language groups differs during sensitive periods of development and affects the perception of speech (Strange & Shafer, 2008). Furthermore, comparison of native-language speech perception and neurophysiological indices of speech processing permits an exploration of the effects of varying auditory-phonetic input from the natural environment on stages of neural speech processing. In addressing these issues, we ask two main questions, (1) What is the role of the acoustic features of phonemes that change with context in speech perception and (2) Does native-language experience modulate earlier levels of auditory processing indexed by sensory-obligatory components of auditory event-related potentials?

Phonotactic patterns of speech (i.e., the permissible sequences of phonemes, including consonant clusters in onset and coda positions) differ across languages (Halicki, 2010; Ewen & van der Hulst, 2001). For example, the English and Polish languages contain the consonant cluster /pt/, but only the Polish language has the cluster in the word onset context (i.e., “ptak”). The exposure during development to the phonotactic patterns of the English or Polish language provides a natural experiment, whereby we can determine the effect of context (word onset) on perception of the /pt/ cluster. The acoustic features of phonemes (or phoneme segments) change with each phoneme context. Therefore, by investigating the effect of context of the /pt/ cluster in the different language groups, we focus on the role of the acoustic features of phonemes in speech perception.

Investigating whether acoustic features are intrinsic to speech perception will contribute to our understanding the development of word recognition. The process by which infants segment words from the speech stream so that word meaning can be learned has not been established. Because the acoustic features of phonemes change with each phoneme context (and stress pattern), the phoneme sequence of each word has a unique acoustic pattern. This unique acoustic word pattern is the phonological information of the native-language input. Infants detect word-units (phoneme sequences) within the first year of life before word meaning is learned (Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk, 1993; Saffran, Aslin & Newport, 1996) indicating that native-language phoneme sequences are encoded in the brain without associated word meaning. After only minutes of exposure to particular 3-syllable nonsense words, eight-month old infants can detect these nonsense words from a string of syllables in which these patterns are embedded (Saffran, Aslin & Newport, 1996). Neural encoding of the unique acoustic word patterns in auditory cortex can explain detection of these specific patterns. Thirteen-month babies can comprehend nonsense words after hearing the words in association with novel objects only nine times, a process termed fast mapping (Woodward, Markman & Fitzsimmons, 1994). This finding suggests that neural encoding of specific speech patterns is efficient during early development. Demonstrating that the context-specific acoustic features of phonemes are encoded at a neural level will supports the view that rapid encoding of unique acoustic patterns of words within auditory cortex is the process supporting word learning early in development and supporting rapid and efficient speech perception of words in the mature system. Furthermore, determining stages for cortical processing of acoustic and linguistic aspects of speech stimuli in adult populations is a necessary for understanding speech perception in adults and during native-language development.

Previous studies have demonstrated that the influence of native-language speech perception begins after approximately 200 ms (Shafer, Schwartz & Martin (in press); Shafer, Schwartz, & Kessler, 2002). The influence of native-language speech perception in the earlier sensory-obligatory components of the ERP waveform from fronto-central (P1N1P2 complex) and lateral temporal sites (T-Complex) has not been investigated. The P2 peak within the P1N1P2 complex has been shown to reflect heightened awareness of auditory stimuli following exposure (Reinke, He, Wang, & Alain, 2003) and, therefore, the P1N1P2 complex has been assumed to precede cortical processes contributing to speech perception in a serial manner (Naatanen &Winkler, 1999; Ceponiene, Alku, Westerfield, Torki & Townsend, 2005; Tremblay & Kraus, 2002). Alternately, it is possible that cortical processing of the acoustic and linguistic aspects of the speech signal involve parallel processes and, therefore, the P1N1P2 complex may not precede cortical processes underlying speech perception (Davis and Zerlin, 1966; Naatanen & Winkler, 1999). The current research examines stages of cortical processing of the acoustic and linguistic aspects of the speech signal in the English and Polish listeners. The information obtained can be used in future research as a basis for comparison with communicatively impaired populations. For example, children having speech and language impairment may show cortical-level processing deficits involving the acoustic and/or linguistic aspects of the speech signal.

In the current project, behavioral and event-related potential measures are used to investigate speech perception and various stages of speech processing. Behavioral tasks of speech perception provide information about the end result of speech perception. Event-related potentials substantiate the behavioral response and provide information regarding earlier stages of speech processing.

Event-related potential (ERP) methodology is a safe, non-invasive time-sensitive measure of neural processing. ERPs provide information about stages of speech processing, which is not available using behavioral methodology alone. ERPs record electrical signals at the surface of the scalp that are time-locked to a stimulus and provide information about underlying neural processes occurring over time in the auditory, cognitive systems. Electrical signals on the surface of the scalp are amplified and the ERP responses to numerous stimulus trials are averaged (Picton, Linden, Hamel & Maru, 1983).

The origin of the cortical ERP is, predominantly, extracellular potentials derived from synaptic activity from large neural ensembles obtained under specific conditions. For example, activity of neuronal groups must be synchronous and neurons within the neural group must have the same spatial orientation. The pyramidal cell, the dominant cell of the cortex, has apical dendrites in parallel allowing for synchronous current flow. Potentials from brain areas with cells that do not have the same spatial orientation (e. g., thalamic relay cells, basal ganglia) are cancelled and will not be detected within the ERP waveform. Hence, ERP waveforms are not a complete representation of neural activity. Also, electrical activity measured at the scalp reflects the summation of neural activity and, therefore, the intracranial source of neural activity cannot be isolated from surface

ERP measures (Steinschneider & Dunn, 2002; Eggermont, 2007). Nevertheless, ERP measures provide information pertaining to stages of neural speech processing that lead to cortical processes underlying speech perception.

The current dissertation examines the effect of a natural variable, exposure or lack of exposure to a particular consonant cluster (/pt/) within a particular context (word onset) during language development on speech perception and on early and late stages of neural speech processing. The methodology permits an investigation of the manner in which the acoustic features of speech are registered in cortical auditory pathways and translated into the integrated neural signal necessary for the process of speech perception.

In chapter 2, we report our investigation of the phonotactic influence on perception of the consonant cluster /pt/ in Polish and English listeners. Specifically, we ask if English listeners perceive the /pt/ consonant cluster in the same manner as native-Polish listeners and whether neural activity, indexed through event-related potentials, matches the perception of the native-language groups. To determine perception, we contrast the onset /pt/ cluster with the segment /pət/ in word onset (e.g., “petunia”) nonsense words. Also, we compare perception of the *pt* contrast (i.e., /pt/ versus /pət/ in word onset within nonsense word stimuli) with an *st* contrast (i.e., /st/ versus /sət/ in word onset within nonsense word stimuli), which is a phonological contrast in both Polish and English. All nonsense word stimuli are potential real words in Polish and English with the exception of nonsense words having onset /pt/ which is an illegal form in English. By comparing two specific sound contrasts, the *st* and *pt* contrasts, that are both phonological contrasts in the Polish language also permits us to examine the effect of the acoustic aspects of native-language phonemes on neural speech processing.

In chapter 3, we investigate whether the acoustic distinctions in the experimental stimuli are registered in the brains of the native-language groups in the same manner. The P1N1P2 complex and the T-Complex, both sensory obligatory components to the physical-acoustic change in ongoing sound stimuli, are examined. Specifically, we ask whether the P1N1P2 and T-Complex responses to the four word forms in the experimental stimuli (i.e., 2-syllable *pt* onset (e.g., /ptimɑ/), 3-syllable *pt* onset (e.g., /pətimɑ/), 2-syllable *st onset* (e.g., /stimɑ/) and 3-syllable *st* onset (e.g., /sətimɑ/) are the same in native-Polish and native-English language groups.

In Chapter 4, we conclude by summarizing the findings in chapters 2 and 3. We discuss the significance of the results and plans for future research.

CHAPTER 2

The phonotactic influence on the perception of a consonant cluster /pt/ by native-English and native-Polish listeners:

A behavioral and ERP study

INTRODUCTION

Sounds within the speech signal are processed by their acoustic features in cortical auditory pathways (Steinschneider, Schroeder, Arezzo & Vaughan, Jr., 1995; Steinschneider, Volkov, Fishman, Oya, Arezzo, & Howard III, 2005). These acoustic features of speech sounds, in some way, are integrated into word-like units to which meaning can be attached. The acoustic features of a phoneme (e.g., /p/) vary depending on the context of that phoneme, that is, the specific sound, which precedes or follows, or the presence or absence of a period of silence preceding or following. As an example, if /p/ is heard in syllable onset (e.g., pat), the phone consists of a rapid *rise* in the first formant frequency during opening of the lips. If /p/, on the other hand, is heard in syllable coda (e.g., tap), the phone consists of a rapid *fall* in the first formant frequency, related to lip closure for the /p/ (Bordon, Harris, & Raphael, 2003 p 113). The phonotactic rules of a language (i.e., the permissible sequences of phonemes, including consonant clusters in onset and coda positions) determine all the allowable contexts of a phoneme. The phonotactic rules of a language, also, determine the stress patterns of a language and the acoustic features of phonemes vary in stressed and unstressed syllables. Stressed syllables, compared to unstressed, have a higher fundamental frequency, increased intensity and duration (Bordon et al., 2003 p 122). By exploring the influence of the phonotactic rules (allowable phoneme contexts) of a language on speech perception, we highlight the role of the acoustic features in auditory speech processing. “In auditory word recognition…it is commonly assumed that some sort of sound-based representation that captures phonological information establishes contact between the acoustic signal and word meaning (Praamstra and Stegeman, 1993, p 74).” We explore the nature of the sound-based representation in word recognition and the role of the acoustic features of phonemes in speech perception.

To explore the sensitivity of the auditory and cognitive systems to the contextual effects of the native-language, we examined the consonant cluster /pt/. The phonotactic rules of English permit the /pt/ consonant cluster to occur at the end of words, but not in word onset. The probability of /pt/ occurring in word onset in English is zero using a phonotactic probability calculator based on a 20,000 word data base (Vitevitch & Luce, 2004). In contrast, the /pt/ cluster occurs at the end of words (i.e., except) and as a regular and irregular past tense form in words (i.e., jumped; slept). In conversational speech, English listeners hear sequences of phonemes such as *ptina*, in sentences like "He sle*pt in a* bed", however, sequences of phonemes like *ptina* are never heard without a preceding vowel. Likewise, the /pt/ cluster is never heard following pauses or following consonants closing preceding words. We, therefore, do not expect native-English listeners to be sensitive to the acoustic features of the /pt/ cluster in word onset.

Phonotactic rules of a language affect production and perception of speech. Consonant clusters, that are not native to a language, are difficult to produce. Lack of exposure to a consonant cluster in the native-language, such as /zg/ in the onset of the nonsense word "zgomu", prevents English speakers from producing the cluster in the same way as native speakers. English speakers produce non-native clusters such as /zg/ without properly timed co-articulation (Davidson and Stone, 2004). Similarly, listeners of a language have difficulty perceiving non-native consonant clusters. For example, because the simple syllable structure of Japanese does not allow for complex consonant clusters, or consonants in the coda (i.e., syllable final position), Japanese listeners perceive a vowel between two consonants when a vowel is not present. A nonsense word such as “ebzo” is perceived as “ebuzo” by Japanese listeners (Dupoux, Hirose, Kakahi, Pallier & Mehler, 1999). Also, Halle, Segui, Frauenfelder and Meunier (1998) found that the legality or illegality of consonant clusters within the same syllable influence perception. Onset clusters, illegal in the French language (i.e., /tl/ and /dl/ within nonsense words, example “tlabdo”), were divided into ten acoustic segments that became progressively longer. The initial segment was approximately 10 ms and the final segment was a whole syllable of approximately 190 ms in duration. Participants perceived the initial segments as dentals (i.e., /t/ or /d/) but once features of the /l/ segment appeared, the clusters were perceived as velar clusters (i.e., /kl/ or /gl/) which are legal phonotactic structures in syllable onset in French. Hence, the clusters were, erroneously, perceived as probable sequences in the native language.

Infants become aware of the phonotactic patterns of the native-language before the end of the first year of life at a time when they are only beginning to comprehend word meanings (Jusczyk, Friederici, Wessels, Svenkerud & Jusczyk, 1993). Jusczyk et al., suggest that infants attend to and encode native-language sound patterns and, subsequently, lexical meaning is attached (p 641). Furthermore, Jusczyk, Luce and Charles-Luce (1994) discovered that 9-month-old infants are not only aware of the phonotactic patterns of their native language, but are aware of the frequency of phonotactic patterns in their language. Nine-month-olds, but not 6-month-old infants, were found to prefer to listen to CVC nonsense word lists that contained phoneme and phoneme sequences that had a high frequency (and high probability) in the language rather than CVC nonsense words containing phoneme and phoneme sequences of low frequency (and low probability) in the language. Together, these findings suggest that neural patterns for word-units begin to form early in development and the frequency of the sound combinations are encoded in the neural pattern. Also, after brief exposure, infants can detect word-units of an artificial language (e.g., 3-syllable nonsense words) from a string of syllables (Saffran, Aslin & Newport, 1996). The researchers suggested that the infants learned the order of syllables within the words through statistical probability learning, an inferred cognitive skill. A cognitive skill may not be necessary, however, because the features of phonemes change with each context making each phoneme sequence unique. Therefore, the acoustic features can provide the link for the phoneme sequence within word-like units. The native-language phoneme sequence is the language experience mapped as the neural code. As the infant matures, the native-language phoneme sequences are processed automatically, whereas, non-native phoneme sequences are not (Strange & Shafer, 2008).

Acoustic features of phonemes that change with context affect lexical activation within priming tasks. (see McQueen, Dahan, Cutler, 2003 for a review of the research). For instance, voice onset time (VOT), has been shown to affect lexical activation within a semantic priming task. A word containing an ambiguous onset consonant (e.g., /k/) with an altered VOT (a sound between a /k/ and /g/) acted as a semantic prime, however, the altered features of the ambiguous sound resulted in slower lexical activation. This pattern suggests that the acoustic features influence lexical processing (Andruski, Blumstein & Burton, 1994). Luce's *Par*adigmatic and *Syn*tagmatic model (PARSYN) model recognizes a sub-phonemic unit, the allophone, in lexical activation. "…Allophonic units are "interconnected" and phonotactic patterns of the language and frequency of sound patterns affect their strength for lexical activation" (Luce, Goldinger, Auer, & Vitavitch, 2000; Vitavitch, 2007). An alternative model is that the acoustic feature is the basic unit of words, which is consistent with cortical speech processing (Steinschneider et al., 1995; 2005). The allophone as a basic unit in word recognition cannot be realized because speech co-articulation does not permit segmentation into allophonic units (Borden et al., 2003). Furthermore, acoustic features as the underlying basic unit of words allows for a feasible word recognition process. Languages have a limited number of sounds and a large variety of words that can be accessed from sound combinations (McQueen et al., 2003). The acoustic features of phonemes in a word's sound sequence constrain the possibilities in word recognition because there is a unique acoustic pattern for each phoneme sequence, which forms the word.

We suggest that recognition of words can be derived directly from acoustic features without an intermediate allophonic or phonemic level of processing. Acoustic features, as the basic unit in word recognition, explain infant detection of phoneme sequences because the acoustic features link the phonemes. Also, Acoustic features as the basic unit, explains the influence of the features in priming studies of lexical activation and constrains the possibilities in word recognition. Furthermore, a model in which acoustic features are the basic unit in word recognition is consistent with cortical speech processing and speech co-articulation.

English listeners experience the /pt/ cluster in their language, but not in word onset, the context examined. The context (i.e., word onset) modifies the acoustic features of the /pt/ cluster. Fine-grained timing of speech processing available from neurophysiological data should reveal sensitivity to acoustic information.

Psychoacoustic Salience

It has been suggested that in addition to native-language input, psychoacoustic salience of phoneme contrasts affects perception (Strange & Shafer, 2008, Burnham, 1986). Burnham (1986) proposed that phonological contrasts vary on a robust/fragile continuum with robust contrasts having a “strong psychoacoustic basis” and fragile contrasts having a “weaker psychoacoustic basis” (p 209). He supported his position by stating that the degree of salience of phonemes’ acoustic features affects how widespread particular phonemes are in the world’s languages, at what age children no longer perceive contrasts when they are not meaningful in the language, and the ease of which the contrasts are restored in older children and adults with training. The two sound contrasts in the current experiment (/pt/ versus /p´t/; /st/ versus /s´t/) differ in the initial phoneme, /p/ or /s/ and, therefore, in their acoustic features (e.g., duration). We compare the two sound contrasts and consider the effect of the acoustic features on speech perception and the relationship to native-language perception.

Salience has, also, been considered in terms of sonority (Berent, Steriade, Lennertz, & Vaknin, 2007). Sonority is defined as “a scalar property of segments correlated with acoustic intensity. Louder segments (e.g., /l/) are more sonorous than quieter segments (e.g., /p/, /t/).” (Berent et al., p 593). Examination of phonotactic structures across the world’s languages suggests that sonority rises from the onset consonant of a word to the vowel and then declines. For example, in “plank” sonority rises from -p- to -l- to -a- and then declines from -a- to -n- to -k. Phoneme sequences within lexical units that follow sonority rules are termed unmarked and are common in the world’s languages and phoneme sequences that do not follow sonority rules are termed marked and are rare. Berent et al., examined the influence of native-language experience on perception of both native and non-native consonant clusters. Furthermore, Berent et al., explored English listeners' sensitivity to the markedness of non-native consonant clusters. Onset consonant clusters were examined in nonsense words that were marked to different degrees on the sonority scale. The nonsense words had either a rise in sonority (e.g., /bn/, /dl/; obstruent-nasal or obstruent-liquid), a plateau (e.g., /bd/, /pt/; obstruent-obstruent) or a fall in sonority (e.g., /mg/, /lb/; liquid-obstruent, nasal-obstruent). If a language has onset consonant clusters with a more marked type of sonority pattern, they will also have the less marked patterns. So if a language has a plateau sonority pattern (e.g., /pt/), then it will also have a small rise sonority pattern (e.g., /pn/) and the expected large rise pattern (e.g., /pl/). Interestingly, the authors point out that the S-stop onset consonants are an exception to the sonority rule and are common in the world's languages. Berent et al., examined these three types of non-native onset consonant clusters in single-syllable nonsense words that became progressively more marked from a small rise in sonority (e.g., "pnaef"), to a plateau in sonority (e.g., "ptaef"), to a fall in sonority (e.g., "rpaef"). The nonsense words were matched with disyllabic nonsense words in which a schwa occurred between the two consonants in the cluster (e.g., "penaef", "petaef", "repaef"). A Russian speaker produced the nonsense word stimuli because the Russian language contains these patterns. A task requesting English and Russian participants to determine the number of syllables in the nonsense words revealed significant group differences. The Russian listeners had no difficulty determining the number of syllables in the stimuli, confirming the importance of the native-language in speech perception. English listeners, on the other hand, performed poorly and their accuracy in identifying the single-syllable nonsense words (e.g., “pnaef”, “ptaef,” “rpaef”) as having one syllable became significantly worse as the consonant cluster types became more marked. The researchers concluded that there is a “knowledge that is related to sonority” but this knowledge does not necessarily have to be sonority but could be some other factor that correlates with sonority (Berent et al., p 624). This research suggests that some characteristic of the consonant clusters examined, none of which were native in the English language, affected perception. In examining perception of the consonant cluster /pt/, we compare native-English and native-Polish listeners because the Polish language has /pt/ in word onset and the English language does not. We compare /pt/ to /st/ because /st/ occurs in both languages in word onset. Also, because the Polish language has both onset clusters, /pt/ and /st/, we consider their differing acoustic features in speech perception.

Event-Related Potential Research

As we have seen, various behavioral methods have been employed in examining the phonotactic influence on speech perception. Behavioral measures provide information about perception or the end result of a task. Event-related potentials (ERP), also employed in speech perception research, provide information about underlying processes in the auditory and cognitive systems during perception of speech and language. Event-related potentials are electrical signals recorded at the surface of the scalp that are time-locked to a stimulus and provide information about underlying neural processes occurring over time in the auditory and cognitive systems. In order to detect the electrical (ERP) response to the stimulus relative to random brain activity, the electrical signals on the surface of the scalp are amplified and the ERP responses to numerous stimulus trials are added and averaged (Picton, Linden, Hamel & Maru, 1983). Electrical activity measured at the scalp reflects the summation of neural activity (Steinschneider & Dunn, 2002).

Acoustic Level of Processing

Evidence of speech processing at the acoustic level has been found for the P1N1P2 complex (Martin, Tremblay & Stapells (2007). The P1N1P2 response to the changing acoustic features or sounds within a word has been termed the acoustic change complex (ACC) (Martin et al., 2007; Kaukoranta, Hari & Lounasmaa, 1987; Ostroff, Martin, and Boothroyd, 1998). The ACC has been shown to be highly sensitive to changes in intensity, frequency and duration (Martin et al., 2007). For example, auditory stimuli having low pitch result in larger N1 and P1 amplitudes than stimuli having high frequency pitch (Crowley & Colrain, 2004); Therefore, words beginning with /p/ are expected to have a larger N1P1 amplitude than words beginning with /s/ phoneme. Auditory cortex on the superior temporal plane is the primary contributor to the P1N1P2 complex recorded from fronto-central electrode sites (Martin, Tremblay, & Korczak, 2008; Steinschneider & Dunn, 2002; Naatanen & Picton, 1987).

The T-complex is a negative-positive-negative wave with the first negative peak around 70 to 80 ms (i.e., Na), a positive peak around 100 ms (Ta) and a large negativity around 140-160 ms (Tb), obtained from lateral temporal electrode sites (Shafer, Schwartz & Martin, in press; Tonnquist-Uhlen, Ponton, Eggermont, Kwong, & Don, 2003; Naatanen & Picton, 1987; Wolpaw & Penry, 1975). Dipole source modeling suggests a T-complex source in secondary auditory cortex on the lateral surface of the superior temporal gyrus and is at least partially independent from the P1N1P2 complex (Ponton, Eggermont, Khosla, Kwong, & Don, 2002).

Evidence of phonological contrast within words/nonsense words (lexical access) has been demonstrated in ERP components beginning around 200 ms post word-onset (Shafer, Schwartz & Martin, in press; Shafer et al., 2002; Connolly & Phillips, 1994). These findings suggest that phonological information is enhanced in the auditory pathways at these later stages of processing (for details pertaining to enhancement of acoustic signals in cortical speech processing see Steinschneider, Fishman & Arezzo, 2003). The acoustic contrasts within the words/nonsense words are likely registered at earlier processing stages. Thus, early components, such as the P1N1P2 complex at fronto-central sites and the T-Complex component (i.e., Na, Ta and Tb components) are expected to reflect acoustic-levels of processing. In contrast, late components, the PMN (Connolly & Phillips, 1994) and the N400 are expected to reflect phonological levels of processing.

Phonological-Lexical Processing

Information about phonological contrasts that are meaningful in the language can be derived from tasks requiring lexical access. The N400 ERP component, a negative deflection around 400 ms post-word onset at centro-medial electrode sites, is elicited during tasks of lexical access. A more negative deflection of the N400 component is interpreted as greater “effort” in accessing a lexical item (Nobre & McCarthy, 1994; McCallum, Farmer, & Pocock, 1984; Nobre & McCarthy, 1994; Holcomb, Grainger, & O’Rourke, 2002). Also, the N400 component has been found to reflect phonological differences between spoken words/nonsense words, with a greater negativity elicited for words and nonsense words that are less similar phonologically than more similar phonologically. The phonological N400 has a somewhat different topography than the "semantic" N400 in that this response was large at lateral parietal sites (Praamstra & Stegeman, 1993; Praamstra et al., 1994; Shafer, Schwartz & Kessler, 2002). Studies by Connolly and colleagues show an ERP component that is sensitive to phonological expectation appearing earlier than the N400, at 275 ms post word onset.

In a recent study, the T-complex and a component peaking around 350 ms from left frontal-temporal sites were examined in response to phonological differences in words (Friedrich et al., 2009). Participants were presented with a phonological priming task in which the prime fragment (1) was identical to the initial syllable of the target word (identity condition, e.g., dra-dragon), (2) was a variation of the initial syllable of the target word in that only the initial phoneme was different (variation condition, e.g., kra-dragon), or (3) was completely different from the initial syllable of the target word (control condition, e.g., hun-dragon). Both the identity and variation conditions showed a reduced T-complex from left lateral electrode sites when compared with the control condition (ISI range: 291 ms - 731 ms). The minimal phonemic contrast distinguishing the identity and variation conditions, however, was not reflected in the T-complex but was reflected in a later positive component peaking around 350 ms from left anterior electrode sites. "Both the T-complex and the P350 suggest a dominance of radial left-temporal neural sources underlying the neural priming effects" (p 5). The less distinctive phonological contrast (meaningful in the language), distinguishing the identity and variation conditions, was not reflected in the earlier T-complex component.

Conscious Responding

Speech perception tasks involve conscious behavior. The late positive component (LPC) has been shown to reflect a conscious process during tasks of stimulus evaluation. The LPC also termed P300, P3, and P3b is a positive wave peaking 300 ms or later that reflects cognitive processing, generally, in tasks requiring an overt response. It is affected by numerous variables involved in stimulus evaluation and response selection. Johnson's P300 amplitude model (1986) claimed that once a stimulus is "transmitted" through factors involving attention and certainty, other variables that affect P300 amplitude can be classified within the categories of stimulus complexity, task complexity, stimulus value and probability. The neural processes reflected by the LPC component are not clearly understood, likely, because it is a composite of multiple brain activity (Picton, 1992; Luck, 2005) and has multiple brain sources (Steinschneider & Dunn, 2002; Picton, 1992).

Eliminating a participants' overt response in behavioral tasks reduces or eliminates the LPC component. For instance, research by O'Rourke & Holcomb (2002) required participants to listen to words and pseudowords in word lists. In a first experiment, participants performed a lexical decision task and an LPC following the N400 component occurred. In a second experiment, participants were instructed "to carefully pay attention to all spoken stimuli" (p 137) but were not required to perform a behavioral response. An LPC did not occur in this follow-up experiment. In addition, Holcomb et al., (2002) found the LPC component to mask the N400 component in a task involving overt responding. In a first experiment, an N400 and LPC component were elicited from participants engaged in a lexical decision tasks using words and pseudowords as stimuli. In a follow–up experiment, participants no longer engaged in a lexical decision task. Rather, they were asked to respond only to “body part” lexical items, which were not included as experimental stimuli. Once again, as in the study by O’Rourke & Holcomb (2002), the LPC was not elicited and the N400 component elicited was more negative than in the first experiment, suggesting that the LPC was masking the N400 response to some degree.

Both perceptual variables involved in stimulus evaluation (e.g., ease of discriminability of stimuli) and variables involved in response selection (e.g., compatibility of response such as a button response performed on the same or opposite side as presented stimuli) have been found to affect the LPC (McCarthy & Donchin, 1981; Leuthold & Sommer, 1998; Luck, 1998; Isreal, Chesney, Wickens & Donchin, 1980). Pertinent to decision making for a native or non-native linguistic contrast is an early finding by Squires, Hillyard and Lindsay (1973) of a positive vertex potential between 300-450 ms (right mastoid reference), termed P3, reflecting the level of certainty in detection of a threshold level tone. When participant’s confidence ratings for detecting a signal tone were high, P3 amplitude was greater than when confidence ratings were low. High confidence ratings for hits (report signal detected and signal occurred) resulted in greater P3 area than low confidence rating for hits. High confidence ratings for false alarms (report signal detected and signal absent) resulted in a greater P3 area than low confidence ratings for false alarms. Also, Picton & Hillyard (1974) demonstrated that the P3 component is affected by attention to the stimulus.

Furthermore, the LPC has been shown to reflect phonological contrast within words (Schwartz, Maxfield, Gross & Shafer, in preparation). In a same-different judgment task, participants heard same and different real word pairs (e.g., caught-caught;

tip-dip; fate-fake) and same and different nonsense word pairs. The words (nonsense words) in the different pairs contained a phonological contrast in word onset or coda and, as in the current study, natural recorded words (or nonsense words) were used as stimuli. An LPC around 500 ms followed an N400 component. The LPC was greater for the second word of the different pairs than for the second word of the same pairs with an earlier latency when the phonological contrast occurred in word onset and a later latency when the phonological contrast occurred in word coda. The LPC response was larger for real word than for nonsense word pairs. In the current experiment, the N400 and the LPC allow us to determine if perception of the sound contrasts in the behavioral task are mirrored in the neural signal (ERP).

Overview

To explore the contextual effects in the native-language, we examined the consonant cluster /pt/ in word onset. Nonsense word stimuli in the current experiment consisted of same word pairs or different pairs. Different word pairs (e.g., /ptimɑ-pətimɑ/) contain a contrast that occurs in English in word final (e.g., “trumped/trumpet”), but not in word onset. In the Polish language, the /pt/ cluster occurs in word onset (e.g., “ptak” – “bird”). The contrast sequence within the nonsense words, that is /pət/ occurs in English (e.g., “petunia”) as well as in Polish (e.g., “petycja” – “petition”; “Petronella” - a fairy tale character). Native-Polish and native-English listeners were tested to determine the effect of native-language experience on perception of the consonant cluster /pt/. Both native-English and native-Polish groups are exposed to the consonant cluster, /pt/ in their language, but only one group (native-Polish) experience the cluster in the phoneme context examined, word onset. The consonant cluster /st/ (e.g., Polish, “stal” – “steel”;

English, “stop”) and the phoneme sequence /sət/ having the unstressed schwa (e.g., Polish, “suterena” – “basement”; English, “sateen”) occur in both languages in word onset and are examined as an experimental control. Furthermore, because both clusters (/pt/ and /st/) occur in the Polish language in word onset, we compare the effect of the acoustic characteristics of the sequences on native-language perception. Our research questions whether English listeners perceive the /pt/ consonant cluster in the same manner as native-Polish listeners and whether neural activity, indexed through event-related potentials (N400 and LPC), matches the perception of the native-language groups. In addition, acoustic measurements corresponding to the /pt/ and /pət/ sequences in the nonsense words were obtained in order to determine physical-acoustic events contrasting the word onset sequences in the stimuli. Also, acoustic measurements compare the /pt/ and /st/ contrasts.

A behavioral task conducted while recording event-related potentials, required participants to determine if the target nonsense word (i.e., second word) in the nonsense word pairs had two or three syllables. The behavioral task, not only required the participant to detect a contrast between the two and three syllable words, but, also required a more difficult meta-linguistic type skill. The level of difficulty of task demands has been found to influence performance in speech perception (Strange & Shafer, 2008). Also research, including training studies, has demonstrated that second language learners have the ability to detect non-native phoneme and consonant cluster contrasts, although the majority of subjects do not achieve a native-like level of performance (Davidson, Shaw & Adams, 2007; Strange & Shafer, 2008; Birdsong, 2005). The task design of determining whether the target nonsense word has two or three syllables was chosen to ensure that participants listened to the whole word, because it was believed that listening to the whole word was necessary for obtaining the N400 component (see Praamstra & Stegerman, 1994). A phonological priming task design was used in the current experiment to examine the presence or absence of a neural response (i.e., ERP) to a phonological contrast. If a linguistically-relevant contrast is detected in the different pairs relative to the same pairs, we expect it to be reflected in ERP components of lexical processing (N400) and conscious processing (LPC).

Hypothesis

Behavioral Task: We hypothesized that Polish listeners would determine whether *pt* nonsense words had two or three syllables with greater accuracy than the native-English listeners. Both groups would perform the task for the *st* nonsense words with the same level of accuracy.

ERP-Target word stimuli: We predicted a larger N400 for the target word in different pairs compared to the target word in same pairs when the phonological contrast was linguistically-relevant for the language group (Praamstra & Stegeman, 1993; Praamstra et al. 1994; Shafer et al., 2002). Therefore, we expected an N400 greater for the *pt* target word in different pairs (i.e., /ptimɑ-*pətim*ɑ/) when contrasted with the *pt* target word in the same pairs (i.e., /pətimɑ-*pətim*ɑ/) for the native-Polish language group, but not the native-English language group, because the contrast /pt/ vs /*pət*/ is linguistically-relevant in word onset only in the Polish language. A greater N400 to the target word of the different pair would indicate new neural firing in response to the novel stimulus word because the contrast that is meaningful in the language was perceived. We expected a greater N400 in response to the *st* target word in the different pairs (i.e.,

/stimɑ-*sətim*ɑ/) when compared to the target word in the same pairs (i.e., /sətimɑ-*sətim*ɑ/) for both language groups because the phonological contrast /st/ vs /sət/ occurs in both English and Polish word onset. The research findings of Schwartz et al., (in preparation) led to an additional prediction. We hypothesized that the phonological contrast within our nonsense word stimuli might also be reflected in the LPC component. Again, the phonological contrast would be reflected in the ERP response, only for the language groups for which the contrast is linguistically relevant. A greater LPC for the linguistically-relevant contrast (in the different pairs relative to the same pairs) would indicate that the contrast segment in the target word was detected within the process of stimulus evaluation necessary for the decision making process. Speech perception of the contrast /pt/ vs /pət/ was predicted to be reflected in the LPC for the native-Polish group, only. Speech perception of the contrast /st/ vs /sət/ was predicted to be reflected in the LPC equally, for both groups. Also, the phonological contrasts may be reflected in the LPC response only, because a large LPC response may mask the N400 component (Holcomb et al., 2002).

If the English listeners are able to detect the non-native language contrast in word onset and this detection is mirrored in the N400/LPC response then the experience of having the sound contrast within the language, albeit within a different context (i.e., /pt/ vs /pət/ contrast in word final context; “trumped/trumpet”) may be sufficient for development of perception of a sound contrast. In contrast, if the English listeners are unable to detect the non-native language contrast in word onset (i.e., /pt/ vs /pət/) within the behavioral task and the performance is mirrored in the N400/LPC response, then the acoustic characteristics of the /pt/ cluster present within the word onset context are encoded within the neural speech signal for speech perception. The stages of neural processing for speech perception are the subject of Chapter 3.

Our predicted findings, if confirmed, will establish the acoustic features of phonemes as intrinsic aspects of the auditory signal in speech processing. We, also, will have demonstrated to what extent the perception revealed in behavioral tasks can be substantiated through neurophysiological means. Furthermore, if the neurophysiological response to the *pt* and *st* contrasts (i.e., /pt/ vs /pət/ and /st/ vs /sət/) differ for the Polish participants, then factors in addition to native-language experience are involved in speech perception.

METHOD

Participants

Thirteen native-English listeners (8 female and 5 male) between the ages of 21 and 35 (mean = 29 years) and 13 native-Polish listeners (8 female and 5 male) between the ages of 23 and 34 years (mean = 30 years) were included in the study. The behavioral data from one male Polish participant was lost due to a computer error. Therefore, behavioral data from 12 native-Polish listeners and 13 native-English listeners was included for analysis. Data from one English participant was excluded from ERP analysis because the participant passed the hearing screening in one ear, only. ERP analysis was, therefore, conducted on 12 English and 12 Polish participants.

The native-Polish listeners were bilingual Polish-English speakers who had emigrated from Poland to the United States as young adults after 15 years of age. The English listeners reported having no exposure to Polish or other Slavic languages that contain /pt/ as a consonant cluster in word onset. Participants were without a history of speech, language, or cognitive impairment. Participants passed a hearing screening at 25dB HL (500 Hz, 1 KHz, 2 KHz, and 4 KHz; Welch-Allyn Audioscope 3, calibrated 1/9/07; All testing was conducted between March & June, 2007). Two Polish and one English participant were left-handed; all others were right-handed.

Stimuli

Natural speech stimuli consisted of nonsense words recorded from a male bilingual Polish-English speaker (age 26 years) who came to the United States at the age of six years old. After learning English, the speaker continued to converse with his family in Polish and, in fact, was not permitted to speak English in the home. The speaker attended a full day of Polish school (through the Polish consulate) once per week through 11th grade.

The nonsense words were recorded with a Shure SM48 microphone, Earthworks microphone, Preamp lab 101 using the software program, Sound Forge, version 4.5 (Build 49) on a Dell Dimension XPS B800 Pentium III computer with Sound Card, Creative SB Live! Basic (WDM). The stimuli were recorded with a sampling rate of 22,050, sample size16 bit, channels mono. The original recordings were saved and transferred to discs. Selected nonsense word stimuli were copied to separate files and preceding and following silence intervals were removed. Selected stimuli were normalized to -20 dBVU and DC offset was set to zero (Sound Forge version 4.5).

Two and 3-syllable nonsense words beginning with /pt/, /pət/, /st/, /sət/ (e.g., /ptukɑ/, /pətimɑ/, /stɛsɑ/, /sətɪlɑ/) were recorded. All stimuli were potential real words in Polish and English with the exception of all nonsense words that began with /pt/, which is an illegal phonotactic form in English. The vowels heard in the penultimate syllable of the nonsense words were vowels that are present in the Polish and English language. Nonsense words in the *pt* and *st* conditions were matched for rhyme as closely as possible. Specifically, the nonsense words in the *pt* conditions had a counterpart matched for rhyme in the *st* conditions (e.g., ptɛsɑ, pətɛsɑ, stɛsɑ, sətɛsɑ) or a counterpart closely matched for rhyme (e.g., ptɪvɑ, pətɪvɑ, stɪfɑ, sətɪfɑ). The mean and range for the nonsense word lengths for each word type are as follows: three-syllable *pt* nonwords, mean = 552 ms, range 481- 675 ms; two-syllable *pt* nonwords, mean = 490 ms, range 417 ms – 604 ms; three-syllable *st* nonwords, mean = 698 ms, range 633 – 801 ms; and two-syllable *st* nonwords, mean = 622 ms, range = 551 – 698 ms.

Four *pt* conditions and four *st* conditions were created, each condition consisting of 100 nonsense word pairs (800 nonsense word pairs in total). The eight stimulus conditions and examples of nonsense word pairs are provided in Table 2.1. Words within pairs were different tokens, but were matched for pitch contour. Word pairs were delivered with a 250 ms inter-stimulus interval (ISI) and a 2000 ms inter-trial interval (ITI).

E-Prime software, version 1.1 (on Dell Precision PWS 670 Intel (R) computer) was used for stimulus presentation, data collection and scoring. Stimuli were delivered in a sound-treated, electrically shielded room and presented through speakers. They were presented at an intensity ranging from 51 to 63 dB SPL (Sound level meter: B&K 2203 with a B&K 4144 mic, “A” weighting).

Procedure

Participants were instructed to decide if the second nonsense word in the pair (i.e., target) had two or three syllables. A response key labeled two on the left side of the keypad was pressed for two syllable words and a response key labeled three on the right side of the keypad was pressed for three syllable words. A key press could be made during presentation of the target word or during the 2-second interval between trials. Participants received verbal instruction and practice on twenty nonsense pairs. Visual feedback specifying correct or incorrect response was provided during the practice portion of the experiment.

Nonsense word pairs for the experimental portion of the design were presented in 10 blocks with 80 pairs in each block. Stimulus pairs were randomized within blocks and the blocks were presented in a random order. A short break was given between each block where the participant was asked how they were faring.

Means scores, median scores, and standard deviation (SD) scores were calculated for the behavioral data. Accuracy on four *pt* and four *st* conditions was calculated and language groups were compared using the nonparametric Fisher-Exact Test (Siegel & Castellan, 1988).

Acoustic Analysis

We questioned why Polish listeners should have poorer performance on the 3-syllable *pt* nonsense words relative to the 2-syllable nonsense words. Therefore, we closely examined the acoustic properties of the stimuli for 9 of our native-Polish participants. Twenty 3-syllable nonsense word stimuli (tokens) were incorrectly identified as having 2-syllables by at least 4 of the subset of Polish participants. These 20 tokens were analyzed using Multispeech waveform and spectrogram tools (Multispeech software on a Dell Intel (R) Pentium (R) computer) and compared to the 243 nonsense word stimuli that were correctly identified as having 2 or 3-syllables by all 9 participants. First, we listened to the initial syllable of the *pt* 3-syllable words for the presence or absence of voicing for the vowel. Note that it is possible to hear voicing in a single syllable that may not be perceived when heard within the whole 3-syllable word. The 243 stimulus nonsense words correctly identified as having 2 or 3-syllables by a subgroup of Polish participants were also analyzed to determine the acoustic features that distinguished the 2 and 3-syllable nonsense words. The prime and target words for all different *pt* pairs were measured for total word length, length of the burst for /p/, length of the vowel and number of pitch periods (i.e., schwa between the /p/ and /t/ in 3-syllable *pt* words), length of the silence and burst for /t/, and length of the penultimate syllable's vowel. Increased length of the penultimate syllable vowel, either for the 2 or 3-syllable nonsense words, would indicate greater stress, a possible acoustic distinction for the Polish listeners. The target word measure (e.g., word length) was subtracted from the *prime* word measure to provide a difference score. For example, the word length of /ptimɑ/ subtracted from the word length for /*pətim*ɑ/ for the nonsense word pair /pətimɑ-*ptim*ɑ/ was obtained. Significance was tested using a single sample T-test.

Additional acoustic measurements compared the duration of the /st/ and /pt/ clusters, duration of the /sət/ and /pət/ segments and duration of the schwa within the word onset segments. A T-test for independent variables was used for statistical analysis.

EEG Data Acquisition

The Geodesic System (Electrical Geodesic, Inc) was used for data collection. A 65 channel Geodesic Sensor Net with silver/silver-chloride (Ag/AgCL) plated electrodes encased in soft sponges was placed on the head. To facilitate conduction, the sponge-encased electrodes were soaked in a saline solution before being placed on the scalp. The placement on the scalp for each electrode was gently rubbed with saline solution and the sponge-encased electrode was carefully positioned. Prior to recording data, we confirmed that impedance levels for each electrode were at or below 40 kOhms, an acceptable level for the high-impedance amplifiers used (Ferree, Luu, Russell & Tucker, 2001). Electrodes placed above and below the eyes monitored eye movements and eye blinks. The EEG was recorded with a sampling rate of 250 Hz, was bandpass filtered between 0.1 and 30 Hz and referenced to Cz.

After recording, the continuous EEG was processed for segmenting, averaging, artifact rejection and baseline correction using Net Station software (Version 4.1.2). The continuous recording was segmented into epochs consisting of a 200 ms baseline and a 1600 ms post-onset segment. Artifact rejection was set at +/- 70 microvolts. Data from bad channels were replaced by spline interpolation. Data was baseline corrected from -100 ms to 0 ms and re-referenced to an average reference.

Data Analysis

Four comparisons were analyzed in the ERP data in response to the target word stimuli. The ERP response to the target word of the same pairs (*pt same, 2-syllable; pt same, 3-syllable; st same, 2-syllable; st same, 3-syllable*) was compared to the ERP response to the target word of the different pairs (*pt different, 2-syllable; pt different, 3-syllable; st different, 2-syllable, st different, 3-syllable*, respectively). For example, the target word of the same pair condition (/*pətim*ɑ-*pətim*ɑ/) was compared to the target word of the different pair condition (/ptimɑ-*pətim*ɑ/); the target word in these two stimulus conditions is the same, the only difference being what precedes the target word, a same word (actually, a different token of the same word) or a different word. A greater response to the target word of the different pair relative to the same would reflect detection of the sound contrast within the different pair.

The four comparisons in the ERP data were examined using Global Field Power (GFP) and current source density maps.

Global Field Power Analysis (GFP)

Temporal regions of interest in the grand mean waveforms were determined using global field power (GFP) to first reduce the data. GFP is the variance across electrode sites and provides an objective reference-free measure of maximum brain activity across latency (Sussman, Steinschneider, Gumenyuk, Grushko, & Lawson, 2008; Shafer, Ponton, Datta, Morr & Schwartz, 2007; Lehmann & Skrandies, 1980). This measure provides an indication of overall power from all electrodes to a response. GFP (for each participant) was calculated by obtaining the standard deviation of the amplitude measure from all 64 channels at each data point across time.

Current Source Density Analysis (CSD)

Current Source Density (CSD) maps were used to identify spatial regions of interest and to predict the source of the current flow in response to target stimuli. CSD is the second spatial derivative of the raw voltage potential (Sussman et. al., 2008; Perrin et al., 1989) performed with Brain Electric Source Analysis (BESA EEG V5.1; 2005). CSD emphasizes shallow depth current flow, which suggests sources in neocortex and CSD transformations correspond more accurately to current sinks and sources within auditory cortex than raw voltage data (Steinschneider et al., 1992; Pernier et al., 1988). Statistical analyses were applied to the CSD transformed data.

ERP data used in the statistical analyses consisted of CSD transformed amplitude values averaged across 48 ms time windows (12 data points). The data are from one electrode site or, in the case of the LPC component, the mean response from two electrode sites. Electrode placements are shown in Figure 2.1. CSD transformed data for the LPC were examined at posterior-parietal electrode sites including the average response from sites 34 and 38 (central-posterior parietal sites), the average response from sites 42 and 46 (right-posterior parietal sites) and the average response from sites 28 and 29 (left-posterior parietal sites). CSD data for a fronto-central positivity at 400 ms

(P400) were obtained from midline-frontal site 4, left-frontal site 9, and right-frontal site 58. CSD transformed data for a temporal component between 280 and 400 ms was examined from temporal electrode sites including, site 20 (left-anterior temporal, AL), site 24 (left-posterior-temporal, PL), site 56 (right-anterior temporal, AR) and site 52 (right-posterior temporal, PR). An early temporal negativity (Tb wave of the T-Complex) peaking at 184 ms was examined at temporal sites including, site 20 (left-anterior temporal, AL), site 24 (left-posterior-temporal, PL), site 56 (right-anterior temporal, AR) and site 52 (right-posterior temporal, PR).

Mixed ANOVA were calculated (12 native-English and 12 native-Polish) as the between subject factor with stimulus, time and site as the within subject factors. Voltage data used in the statistical analyses consisted of amplitude averaged across 48 ms time windows (12 data points). Only significant main effects and interactions involving condition are reported. Significant interactions were followed up with a step-down analysis or with Tukey's HSD post hoc tests ($p < .05$). Greenhouse-Geisser correction was applied for three or more factors of site. The Fisher-Exact Test, similar to Chi Square, but for small samples, (Siegel & Castellan, 1988) was used to examine differences in behavioral accuracy between the groups.

RESULTS

Behavioral Results

st conditions: Mean, median and standard deviation scores for the four *st* conditions can be seen in Table 2.2 and Figure 2.2. Both the Polish and English groups determined whether the target word in the *st* pairs had 2 or 3 syllables with near perfect accuracy. Accurate performance on the syllable identification tasks indicates that the participants were able to distinguish the 2 and 3 syllable *st* words. Twelve Polish (out of 12 participants) and 11 English participants (out of 13 participants) responded with 90% accuracy or greater for all four conditions. Using the 90% criteria, the Fisher-Exact Test revealed the language groups were not different (p = .4800).

pt conditions: Mean, median and standard deviation scores for the four *pt* conditions can be seen in Table 2.2, Figure 2.3. The native-Polish listeners determined whether the target words in the *pt* conditions have 2 or 3 syllables at well-above chance levels. The English listeners responded at above-chance levels for the 3-syllable *pt* condition target words, but only at chance level for the 2-syllable target words. Recall that the English language has 3-syllable words in word onset (e.g., petunia) but does not have 2-syllable words (/pt/ onset). Using a criteria of 80% or greater on all four *pt* conditions for each individual, The Fisher-Exact Test revealed a significant group difference (p = .0052). The Polish participants scored high on the *pt* conditions with mean scores of 90% and 94% on the 2-syllable target word conditions and 81% and 84% on the 3-syllable target word conditions. Mean scores, for the English group, were similar to the Polish group on the 3-syllable target word conditions (i.e., 85% and 82%). Mean scores on the 2-syllable target conditions were low for the English group, 43% and

42% on the same 2-syllable and different 2-syllable conditions, respectively. These findings indicate that the English participants identified the 2-syllable *pt* words as having two syllables at chance levels (see individual performance described below).

Individual data, st conditions: Percentage correct scores for individual participants for the *st* conditions showed high accuracy on all conditions for both groups with little variability between subject scores as shown in Table 2.3.

Individual data, pt conditions: Examination of the individual data for the *pt* conditions revealed that the Polish participants were consistent in their pattern of response. In contrast, the English participants showed different patterns of responding (Table 2.4). The majority of English participants scored well above chance levels for the 3-syllable target conditions, but scored from 0% to 50% correct response when presented with 2-syllable target nonsense words. This indicates that some English participants, when presented with 100 two-syllable nonsense words beginning with /pt/ incorrectly responded 100 times indicating that the words had 3-syllables. A few English participants responded with very high accuracy levels for the 2-syllable target nonsense words. When these same subjects were presented with 3-syllable target nonsense words, however, responding fell to chance levels indicating that the contrast was not perceived and that they showed a bias for the 2-syllable response. As shown in Table 2.4 and Figure 2.3, variability of scores was high for the English participants on the *pt* conditions, as expected.

Summary pt and st stimuli: In summary, both the English and Polish listeners identified the *st* nonsense words as having 2 or 3-syllable with a high degree of accuracy indicating that they differentiated the 2 and 3-syllable *st* word forms with ease. The

Polish group identified the 2 and 3-syllable *pt* words with better than 80% accurately and with a consistent pattern of response. In contrast, the English listeners' performance on the syllable identification task demonstrated that they were unable to distinguish the 2 and 3-syllable *pt* words. The English group identified the 2-syllable *pt* words at chance levels. In addition, individual English participants who identified the 2-syllable *pt* words as having two syllables at better than chance levels, identified the 3-syllable words at chance levels only indicating that they were unable to distinguish the 2 and 3-syllable *pt* words. Only one out of the thirteen English participants was able to identify both the 2 and 3-syllable *pt* words at better than chance levels (see Table 2.4, English participant 5). This English participant reported being fluent in the German language, possibly influencing speech perception of the *pt* contrast.

Acoustic analysis of *pt* conditions

Statistical analysis of duration measurements that compare the mean duration for the *st* and *pt* experimental stimuli follows. The mean duration of the onset /st/ and /pt/ consonant clusters in our experimental stimuli differed by 97 ms (mean /pt/ cluster: 135 ms, SD 11.6 ms; mean /st/ cluster 232 ms, SD 23.5 ms: t = - 37.469, p < .01). The mean duration of the onset /sət/ and /pət/ segments differed by 108 ms (mean /pət/ segment: 173 ms, SD 13.2; mean /sət/ segment: 281 ms, SD 24: t-value 34.6, p < .01). The schwa within the /sət/ and /pət/ segments also differed. The mean duration for the schwa within the /pət/ segment was 27 ms (SD 13.4) and the mean duration of the schwa within /sət/ segment was 36 ms (SD = 11.8) (T = - 4.440; p < .01).

The following results include only *pt* stimuli within correct response trials (for Polish participants). Total word length was greater for 3-syllable *pt* nonsense words than for the 2-syllable nonsense words ($p < .01$; mean difference for two contrast conditions were 60 and 61, respectively. The presence of a vowel and pitch periods (i.e., schwa between the /p/ and /t/ consonants) was also different for the contrast nonsense pairs. The 3-syllable *pt* nonsense words had visible evidence of a vowel with a mean length of 37 ms (range: 18 – 72 ms; $p < .0005$). One to 5 pitch periods were counted for the schwa between the /p/ and /t/ consonants in the 3-syllable nonsense words. Length of the burst of the /p/, length of the silence for occlusion of /t/, length of the burst for /t/, and the length of the vowel in the penultimate syllable were not different for the 2 and 3-syllable nonsense words.

Twenty 3-syllable nonsense word stimuli (tokens) were incorrectly identified as having 2-syllables by at least four of nine Polish participants, a subset of the 12 Polish participants in the experiment. The acoustic characteristics of these 20 tokens were compared to the 243 nonsense word stimuli that were correctly identified as having 2 and 3-syllables by all 9 participants. First, we listened to the initial syllable of the *pt* 3-syllable words for the presence or absence of voicing for the vowel. Voicing was not apparent on the vowel for the 20 tokens that were incorrectly identified (3-syllable *pt* items). In contrast, voicing of the vowel in the initial syllable was perceived within all 3-syllable words within the 243 trials having correct responses (i.e., 111 3-syllable *pt* words). This procedure indicated that there were 20 nonsense word stimuli that may not have been spoken in a native-like manner by our bilingual English-Polish speaker accounting for the incorrect responses to the 3-syllable items by the native-Polish group. It is also possible that some part of the poorer performance of the native-Polish listeners perceiving 3-syllable *pt* nonsense words (compared to 2-syllable *pt* nonsense words) may be due to features related to sonority and, therefore, the 3-syllable *pt* words may be inherently more difficult for the Polish listeners (see Berent et al., 2007).

Because these twenty 3-syllable *pt* stimuli incorrectly produced by our speaker were identified, a subset of ERP data was created by removing experimental trials containing the twenty incorrectly perceived token nonsense words from the data set along with the matched 2-syllable *pt* nonsense words and 2 and 3-syllable *st* nonsense words. (see Appendix A for details and the ERP response to the data subset shown in Figures A.2.1 through A.2.4)

ERP responses to target word stimuli

The results for the target word data are presented in the order of late latency ERP components to early latency components. GFP for the grand mean data from each language group revealed latency regions of interest. Spatial regions for the latencies of interest were then determined from grand mean current source density maps. Temporal sites were examined because previous literature suggests language-related differences reflected from lateral temporal sites (Friedrich et al., 2009). Also secondary auditory cortical activity on the lateral surface of the superior temporal gyrus has a small reflection area on the surface of the scalp and may not be evident in the GFP. In this manner, the following components were identified for analysis: 1. A late positive component (LPC) 2. A positivity around 400 ms (P400) 3. A temporal negativity (280 ms, 328 ms, 376 ms; TN) 4. Tb component of the T-complex, (184 ms).

Late latency ERP response

GFP showed differences between the same and different targets in a late time interval for both language groups in response to the *st* contrast and, for the Polish group only, in response to the *pt* contrast. GFP and CSD analysis follow for the *st* and *pt*-contrasts in the 3 and 2-syllable conditions. Also, difference waves (i.e., ERP response to the target word of the different pair minus the response to the target word of the same pair) were calculated. The ERP responses to the same and different conditions were compared using ANOVAS for the 2-syllable and 3-syllable conditions, separately. Recall that the 3-syllable different pair contains a 3-syllable target word preceded by a 2-syllable word and the 2-syllable different pair contains a 2-syllable word preceded by a 3-syllable word.

Late Latency Cortical ERP Responses: *st* contrast

GFP st contrast 3-syllable: GFP analysis revealed apparent differences in the time windows between 472 ms and 712 ms, as shown in Figure 2.4. Analysis of the 6 time intervals between 472 ms and 712 ms, revealed a main effect of condition (Condition (2) x Time (2) x group ANOVA; F (1,22) = 3.662; p = .006; partial eta squared = .294) and no significant interactions involving condition and language group. The response to the different target was greater than the same target by .424 µV (mean different = 2.497 µV, std.err. = .193; mean same = 2.073 µV, std. err. = .144). A language group difference does not quite reach significance (F (1, 22) = 3.662, p = .069). The English group showed a larger difference between the same and different targets than the Polish group. An additional unpredicted pattern was a significant latency x group interaction (F (5, 110) = 4.808; p = .001; partial eta squared = .179) with a rise in GFP at later latencies for the English participants and a fall in GFP at later latencies for the Polish participants.

GFP st contrast 2-syllable: A peak in GFP is evident for the 2-syllable *st*-target words in the same and different pairs for both language groups beginning around 372 ms and continuing through approximately 856 ms, as shown in Figure 2.5. A very large difference for the contrast is evident for the English group relative to the Polish group. The main effect of condition (472 ms – 712 ms: F (1,22) =4.058; p = .056; partial eta squared = .156) approached significance and the interaction of condition by language group was significant (F (1,22) =4.382; p = .048; partial eta squared = .166). The response to the different target was greater than the same target by .442 µV (mean different = 2.528, std.err. = .259; mean same = 2.086 uV, std. err. = .128).

CSD LPC: *st* contrast

The late latency GFP corresponded to a posterior parietal positivity in the CSD maps as shown in Figure 2.6.

CSD LPC st contrast 3-syllable: An analysis, comparing 2 conditions (same, different) X 7 time intervals (424 through 712ms) X 3 sites (midline, right, left) in all participants (12 Polish, 12 English) revealed a main effect of condition (F (1,22) =13.643; p = .001; partial eta squared = .383). The response to the different target was greater than the same target by .051 µV (mean different = .111, std.err. = .012; mean same = .060, std. err. = .012). A significant interaction involving condition x site x group (F (2, 44) =3.241; p = .049; partial eta squared = .128) is shown in Figure 2.7. A step-down procedure followed for each electrode site, separately. Analysis at averaged midline posterior parietal sites (34/36) revealed a main effect of condition (F (1,22) =4.607; p = .043; partial eta square = .173) and no significant interactions. The response to the different target was greater than the same target by .047 µV (mean different = .098, std.err. = .025; mean same = .051, std. err. = .024). Analysis at the averaged right-posterior parietal site (42/46) found a significant main effect of condition (F (1, 22) =12.525; p = .002; partial eta square = .363) and a significant interaction of condition x group (F (1, 22) = 6.421; p = .019; partial eta square = .226). Tukey HSD applied to the condition x group interaction revealed only the English participants' response to the same and different target to differ significantly (p < .01) at the right-posterior parietal site as shown in Figure 2-8. The response to the different target was greater than the same target by .067 µV (mean different = .113, std.err. = .016; mean same = .046, std. err. = .018). Analysis at the left-posterior parietal sites (28/29) in response to the *st* contrast (3-syllable) revealed no main effect of condition or interaction.

CSD LPC st contrast 2-syllable: A 2 (same, different) X 7 (424ms – 712 ms) X 3 sites (midline, right, left) ANOVA for the English and Polish participants revealed a main effect of condition (F (1, 22) =7.117; p = .014; partial eta square = .244). A greater positive response for the target word preceded by a different word than the same occurred with a difference of .033 µV (mean different = .087, std.err. = .013; mean same = .054, std. err. = .016). A significant interaction of condition x time (F (6,132) = 3.918; p = .021; partial eta squared = .151) was evident and post hoc comparisons found the *st* contrast to be significantly different for the four time intervals between 472 and 616 ms (p < .01). A significant condition x site (F (2, 44) = 6.487; p = .004; partial eta squared = .228) and follow-up Tukey HSD found the *st* contrast (2-syllables) to be different only at averaged left-posterior parietal sites (28/29) (p < .01).

Summary, Late Latency ERP Response st conditions: In summary, analysis of GFP in response to the *st* conditions revealed a late latency component with a peak beginning around 500 ms. The CSD analysis indicated that the GFP difference was greatest at posterior parietal scalp sites within the LPC component. Detection of the *st* contrast was reflected within the LPC by both language groups from midline and right-posterior parietal sites. Furthermore, a greater LPC response to the *st* contrast for the English group relative to the Polish group was evident at the right-central posterior parietal sites (3-syllable). Also, 2 and 3-syllable conditions showed a different topography, with an LPC difference for the 2-syllable contrast significant only from left-posterior parietal sites. A rise in GFP at later time intervals for the English participants and a fall in GFP at later time intervals for the Polish participants (3-syllable) was significant suggesting processing differences for the language groups.

Late Latency Cortical ERP Responses: *pt* contrast

GFP pt contrast 3-syllable: The response to the *pt*-target word (3-syllable) in the different pair resulted in a larger GFP than in the same pairs for the Polish language group but not the English group beginning around 328 ms and continuing through 712 ms, as shown in Figure 2.9. A sharp peak at 400 ms and a broader flat peak after 500 ms are apparent. A 2 (Condition) x 8 (Time) x 2 group ANOVA revealed a latency x group interaction similar to that found in response to the *st* stimuli F (7, 154) = 11.680; p = .000; partial eta squared = .347). No significant main effect or interaction including group was found (Condition x group: F (1, 22) = 1.728; p = .202) for the Polish and English participants between 376 and 712 ms.

GFP pt contrast 2-syllable: A larger GFP for the 2-syllable *pt* words in the different pairs occurs for the Polish group but not the English language group beginning around 300 ms as shown in Figure 2.10. A greater GFP response to the contrast occurs for a long time window beginning around 400 ms, reflecting the LPC component. A 2 (Condition) x 8 (Time) x 2 group ANOVA revealed a significant interaction of latency x group, which does not interact with condition (F (7, 154) = 10.276; p =.000; partial eta squared = .318). The apparent differences to the 2-syllable *pt* target words in the same and different conditions did not reach significance (Condition x group: F (1, 22) = 1.995; p =.172) for the combined Polish and English participants.

CSD LPC: *pt* contrast

The late latency GFP in response to the *pt* contrast corresponded with a posterior parietal positivity in the CSD maps as shown in Figure 2.11.

CSD LPC pt contrast 3-syllable: Analysis of the 3-syllable *pt*-contrast in a 2 (same, different) X 8 (376 ms through 712 ms) X 3 (midline, right, and left) x 2 (Polish, English) mixed ANOVA found no main effect of condition but an interaction of condition by group that was marginally significant (F (1, 22) = 3.96; p= .059; partial eta squared = .145). A step down procedure, analyzing the language groups separately revealed no main effect of condition or significant interactions for either group. As shown in Figure 2-8, English and Polish language groups demonstrated a different pattern of mean response to the *pt* contrast (3-syllable). Whereas, Polish participants showed a larger mean response to the different target, English participants showed a larger mean response to the same target at right posterior parietal sites. The difference wave analysis (below), therefore, compares the group response.

CSD LPC pt contrast 2-syllable: A Condition (2) x 8 (Time) x 3 (Site) x 2 groups mixed ANOVA found a main effect of condition (F (1, 22) = 6.086; p= .022; partial eta squared = .217). The LPC to the different target was greater than the same target by .025

μV (mean different = .049, std.err. = .010; mean same = .024, std. err. = .011). No significant interactions were found. The difference wave below compared the group response.

CSD LPC Difference Wave pt contrast 3-syllable: A Condition (2) x 8 (Time) x 3 (Site) x 2 groups mixed ANOVA for the *pt* difference wave showed a marginally significant main effect of group (F (1,22) =3.737; p = .059; partial eta squared = .145) and no significant interactions.

CSD LPC Difference Wave pt contrast 2-syllable: A Condition (2) x 7 (Time) x 3 (Site) x 2 groups mixed ANOVA did not reveal a significant main effect or significant interactions.

Summary, Late Latency ERP Response pt conditions: The Polish group showed a greater GFP to the *pt* different target relative to the same and the English group showed no effect or a reverse effect. The apparent GFP differences between the English and Polish groups, however, did not reach significance. CSD analysis revealed that the GFP difference was found at posterior parietal sites. Statistical analysis of the response within the LPC component revealed language group differences approached significance. Also, a different processing pattern, not related to the sound contrast, for the language groups is evident in GFP in response to the *pt* conditions (3 and 2-syllable).

Comparison of the st 3-syllable and pt 3-syllable contrast in difference wave: A 2 (st difference wave, pt difference wave) x time (7) x site (3) x group mixed ANOVA found a main effect of condition (F (1, 22) = 8.545; p = .008; partial eta squared = .28) and an interaction of condition x group that approached significance (F (1, 22) = 4.013; p = .058; partial eta squared = .154). Tukey's post hoc comparison revealed that English *st* and *pt* difference waves were significantly different from each other (p = .011), whereas, the Polish *st* and *pt* difference waves were not (p = .914). The main effect of condition for the combined groups found the *st* wave and *pt* wave to differ by .109 µV (mean different = .114, std.err. = .012; mean same = .005, std. err. = .008). Analyzing the *st* and *pt* difference waves (3-syllable) in the English group, separately, found a significant main effect (F (1, 11) = 10.282; p = .008; partial eta squared = .483; *st* and *pt* waves differed by .076 µV (st mean = .066 std err. = 020; pt mean = -.01, std. err = .009). Analysis of the Polish group found no significant effects.

This comparison established that the Polish group response to the *pt* contrast was significantly different than the English response because the *st* difference wave was significantly different from the *pt* difference wave only for the English group.

Cortical ERP responses between 232 ms and 424 ms: *st* contrast

GFP analyzed in response to the *st and pt* contrasts between 232 ms and 424 ms, captures a middle peak in the GFP waveform as shown in Figures 2.4, 2.5, 2.9, and 2.10. The middle GFP peak corresponded to a fronto-central positivity and a temporal negativity evident in CSD maps as shown in Figures 2.12 through 2.15.

GFP st contrast 3-syllable: Analysis of 5 latency periods, 232 ms through 424 ms revealed a significant main effect of condition (F (1, 22) =15.571; p = .001; partial eta squared = .414 and a significant interaction of condition x time (F (4, 88) =4.149; p = .004; partial eta squared = .159. Post hoc comparisons found the contrast to be significantly different for the last four time intervals between 280 and 424 ms (p < .01).

GFP st contrast 2-syllable: As shown in Figure 2.5, the pattern of greater response to the different target relative to the same in response to the 2-syllable condition begins at 376 ms. Analysis of the time intervals 376 and 424 ms revealed a significant main effect of condition (F (1, 22) = 12. 648, p = .002, partial eta squared = .365) with the different target showing greater amplitude than the same target by .495 μV (mean different = 2.272, std.err. = .136; mean same = 1.777, std. err. = .108). A condition x time x group interaction was also significant (F (1, 22) = 4.954; p = .037; partial eta squared =.184). A step down analysis of the English group revealed a significant main effect of condition (F (1,11) = 9.929; p = .009; partial eta squared = .474) and condition x time interaction (F (1,11) = 5.639; p = .037; partial eta squared = .339). The *st* contrast was significantly different at both time intervals (Tukey HSD: p < .01). Analysis of the Polish group found no significant main effect or interaction.

CSD: A peak in GFP between 232 and 424 ms to the *st* contrast, 3-syllable conditions, and at a later time window for the 2-syllable conditions, corresponded to a fronto-central positivity and a left-posterior temporal negativity in CSD maps as shown in Figures 2.12 and 2.14. The fronto-central positivity, which peaks at 400 ms will, henceforth, be termed P400 and the temporal negativity will, henceforth, be termed TN. An inverse relationship is evident for the P400 and TN waveform response as shown in Figure 2.16.

CSD P400: *st* contrasts

CSD P400 st contrast 3-syllable: A (2 (same, different) x 2 (376, 424 ms) x site (central, left, right) x group ANOVA revealed no main effect of condition and a significant interaction of condition x site (F (2, 44) = 4.897; p = .015; partial eta squared = .176). A large positive amplitude was found for the different compared to the same target at the frontal-central site 4 (p < .01). Analysis at fronto-central site (site 4) found a main effect of condition (F (1,22) = 6.567; p = .018; partial eta squared = .230) and no significant interactions involving condition. The different target was more positive to the same target by .102 µV (mean different = .144, std.err. = .043; mean same =.042, std. err. = .024).

CSD P400 st contrast 2-syllable: The response to the 2-syllable *st* contrast revealed no significant main effect of condition or interaction.

Cortical ERP responses between 232 ms and 424 ms: *pt* contrast

GFP pt contrast 3-syllable: A greater GFP difference between the same and different targets to the *pt* contrast (3-syllable) was evident for the Polish compared to the English group for the time intervals between 232 and 424 ms. The condition (2) x time (5) x group (2) ANOVA revealed a main effect of condition that approached significance (F (1, 22) = 3.833; p = .063; partial eta-squared = .148). A condition (2) x time (5) ANOVA on the Polish group separately revealed a main effect of condition (F (1, 11) = 6.565; p = .026; partial eta-squared = .374). The conditions differed by .238 µV (mean different target = 1.995, std. err. = .165; mean same target = 1.757, std. err. = .152). The same analysis on the English group data separately revealed no significant main effect (p = .82) or interactions.

GFP pt contrast 2-syllable: The same analysis for the 2-syllable condition found a significant condition x time x group interaction (F (4, 88) = 3.207; p = .017; partial eta-squared = .127) showing a larger mean GFP to the different target relative to the same for the final two time intervals for the Polish group data but not the English data as shown in Figure 2.17. Analysis of the English and Polish groups separately, in a step down procedure, found no significant main effect or interaction.

CSD P400: *pt* contrasts

The peak in GFP around 375 ms to the *pt* contrast (3-syllable) corresponds to a fronto-central positivity and temporal negativity in CSD maps as shown in Figure 2.13 and 2.15.

CSD P400 pt contrast 3-syllable: A condition (2) x time (376, 424 ms) x site (central, left, right) x group ANOVA in response to the *pt* contrast (3-syllable) revealed a main effect of condition (F (1, 22) = 9.297; p = .006; partial eta squared = .297). The difference between the different target and the same target was .031 µV (mean different = .053; std. err. = .021; mean same = -0.011; std. err. = .02). A significant interaction of condition x time x site x group was also found (F (2, 44) = 4.413; p = .037; partial eta squared = .167). A step down procedures analyses site. An analysis of the midline fronto-central site revealed a main effect of condition (F (1, 22) = 5.372; p = .030; partial eta squared = .196) and a significant condition x time x group interaction (F (1, 22) = 4.285; p = .05; partial eta squared = .163). The targets differed by .063 µV (mean different = .062, std. err. = .032; mean same = -.001, std. err. .029. Further step down analysis of the response to the *pt* contrast from the fronto-central site at 376 ms in a condition (2) x group (2) ANOVA revealed a significant main effect of condition (F (1, 22) = 6.174; p = .02; partial eta squared = .219). The means differed by .096 µV (mean different = .082; std. err. = .032; mean same = -.014, std err. = .03). The fronto-central site examined at 424 ms in a condition (2) x group ANOVA revealed a significant main effect of condition (F (1, 22) = 4.303; p = .05; partial eta squared = .164). The means differed by .058 µV (mean different = .043; std. err. = .033; mean same = -.015, std err. = .029). A condition (2) x time (376, 424 ms) x group ANOVA from the right-frontal site revealed a main effect of condition (F (1, 22) = 7.603; p = .011; partial eta squared = .257). Analysis of left-frontal site for the latencies 376 and 424 ms revealed no significant main effect or interaction. Fronto-central and right-frontal sites of interest are examined in each group separately. The English group (Condition (2) x Time (2) x Site (2)) demonstrated a significant main effect of condition (F (1, 11) = 9.467; p = .011; partial eta squared = .463). The targets in the contrast conditions differed by .078 µV (different mean = .072, std. err. = .026; same mean = -0. 006; std. err. = .026). The Polish group showed no significant main effect of condition and a condition x time interaction that approached significance. (F (1, 11) = 4.193; p = .065; partial eta squared = .276; Tukey HSD: 376 ms p = .000; 424, p = .019).

CSD P400 Difference wave pt contrast 3-syllable: A time (376, 424 ms) x site (central, left, right) x group ANOVA found a significant time x site x group interaction (F (2, 44) = 4.413; p = .037; partial eta squared = .167). A step-down procedure analyses site. As shown in Figure 2.18, analysis of the midline fronto-central site revealed a significant time x group interaction (F (1, 22) = 4.285; p = .05; partial eta squared = .163) in which the Polish group showed an earlier positivity at 376 ms relative to the English group. Post hoc analysis found the group response was not significantly different at either time interval. Analysis of left and right frontal sites separately revealed no main effect or interaction.

CSD P400 pt contrast 2-syllable: A condition x time (376, 424 ms) x site (midline, left, right) ANOVA revealed no main effect of condition or interactions.

CSD P400 Difference wave pt contrast 2-syllable: A comparison of Time (2) X Site (3) x group revealed no main effect of group and a significant interaction of time x group (F (1, 22) = 16.961; p = .000; partial eta squared = .435). Tukey HSD comparisons revealed the language groups were not significantly different at either latency.

Summary, P400: In summary, neural detection of the acoustic difference between the *pt* and *st* different pairs by both language groups was reflected in the P400 response. Different group processing latencies in response to the *pt* contrast (3-syllable), however, suggests influence of the native-language. Analysis of the difference wave (*pt* 3-syllable) at the midline fronto-central site revealed a significant group x time interaction showing an earlier latency response to the *pt* contrast by the Polish participants. Post hoc comparisons of the group response at each time interval were not significant.

CSD TN: *st* contrast

CSD TN st contrast 3-syllable: A condition (2) x time (328, 376) x site (AL, AR, PL, PR) x group ANOVA revealed no main effect of condition, a significant condition x time (F (1,22) = 5. 788; p = .025; partial eta squared = .208) and condition x site interaction (F (3, 66) = 3.053 p = .048; partial eta squared = .122). Response to the contrast in the same and different condition was significant at both time intervals (Tukey HSD: p < .01). Furthermore, post hoc comparisons of the condition x site interaction revealed the response to the *st* contrast to be significant only from the left-posterior temporal site (p = .022). The response did not interact with group. Analysis of language groups separately did not reveal significant main effects of condition or interactions. A step-down procedure to explore the hemisphere effect analyzed the posterior temporal sites, separately. A condition (2) x time (328, 376) x site (PL, PR) analysis revealed no main effect of condition, a significant condition by time (F (1,22) = 5. 204; p = .033; partial eta squared = .191) and condition x site interaction (F (1,22) = 9.728, p = .005;

partial eta squared = .307. Post hoc comparisons found the response to the contrast to be significant at both time intervals 328 ms (p = .000) and 376 ms (p = .015), but only from the left-posterior parietal site (p = .009) as shown in Figure 2.19. Analysis of the hemisphere effect for the anterior temporal sites found no significant main effect or interactions.

CSD TN difference wave st contrast 3-syllable: A time (2) x site (4) ANOVA revealed a main effect of time with the difference wave being more negative at 328 ms than 376 ms. (F (1, 22) = 5. 788; p = .025; partial eta squared = .208). Analysis of the 328 ms period separately finds no main effect.

CSD TN st contrast 2-syllable: Analysis of the 2-syllable condition revealed no main effect or significant interaction.

Summary, CSD TN st conditions: In summary, CSD analysis revealed a left-posterior temporal negativity in response to the 3-syllable *st* different target (relative to the same) that did not differ for the language groups. The conditions differed significantly at both time intervals, but the effect was greater at 328 ms.

CSD TN: *pt* contrast

CSD TN pt contrast 3-syllable: A condition (2) x time (328, 376 ms) x site (AL, AR, PL, PR) x group ANOVA revealed no main effect of condition and a significant condition x time interaction (F (1, 22) = 8. 783; p = .007; partial eta squared = .285). Tukey HSD post hoc analysis found the response to the *pt* contrast to be significantly different at both time intervals 328 ms and 376 ms (p < .01) although larger at 328 ms. Condition and group did not interact (p = .65), indicating the English group response was not different from the Polish. The analysis from the four (AL, AR, PL, PR) sites applied to the Polish group separately, however, found a significant condition x time interaction (F (1, 11) = 5.516; p = .039; partial eta squared = .334) that was not found in the English group data. The Polish group data showed a larger effect of condition at 328 ms than 376 ms. Post hoc comparisons of the condition x time interaction, however, found the *pt* contrast (3-syllable) response to differ significantly at both latencies (328 ms, p = .000; 376 ms, p = .015). Also, a greater left-posterior temporal response was not evident for the *pt* contrast using a combined group analysis or an analysis with the language groups examined separately, as shown in Figure 2.19.

A step down analysis of site follows, which more carefully examines the condition x time interaction. Analysis of data from the left-posterior temporal site in a 2 (Condition) x 2 (328, 376) x group ANOVA revealed no main effect and a significant interaction of condition x time (F (1, 22) = 8.782; p = .007; partial eta squared = .285). Post hoc comparisons revealed the conditions to differ significantly only for the first time interval, 328 ms (p < .01) as shown in Figure 2.20. The same analysis of data from the right-posterior temporal site revealed no main effect of condition or interactions. Analysis of the right-anterior temporal sites in a condition (2) x time (328, 376 ms) x group ANOVA revealed a main effect of condition (F (1, 22) = 4.992; p = .036; partial eta squared = .185). The conditions differed by -.078 μV (mean different = -.039, std. err. = .035; mean same = .039, std. err. = .021). Also, a significant interaction of condition x time (F (1, 22) = 4.818; p = .039; partial eta squared = .180) was found. Post hoc analysis of the condition x time interaction revealed the conditions to differ significantly for both time intervals (p < .01). Analysis of the left-anterior temporal site revealed no main effect or significant interactions.

The right and left-anterior temporal sites each analyzed separately in the Polish group revealed no main effects or interactions involving condition. Also, the right and left-anterior temporal sites each analyzed separately in the English group revealed no main effects or interactions involving condition.

The TN was also analyzed at earlier time intervals between 280 and 328 ms because averaged raw voltage data waveforms as shown Figure 2.21 revealed an earlier left temporal negative wave for the *pt* contrast relative to the *st* contrast. A condition (2) x time (280, 328 ms) x site (AL AR, PL, PR) x group ANOVA revealed no significant main effect of condition or significant interactions. Also, the possibility of a more anterior negativity than posterior negativity for the *pt* contrast was explored. A condition (2) x time (280, 328 ms) x site (AL, AR,) x group ANOVA revealed a main effect of condition that approached significance (F (1,22) = 4.105; p = .055; partial eta squared = .157) and no significant interactions. The analysis for the right and left posterior temporal sites at 280 and 328 ms revealed no significant main effect or interaction. Analysis of the anterior and posterior temporal sites separately for the 280 ms time interval in 2 (Condition) x 2 (Sites) ANOVAs revealed no main effect of condition or significant interactions involving condition.

CSD TN pt contrast 2-syllable: A condition (2) x time (328 ms, 376 ms) x site (AL, AR, PL, PR) ANOVA revealed no main effect of condition and a significant condition x time x site x group interaction (F (3,66) = 2.783; p = .05; partial eta squared = .114). A step down analysis of the Polish group in a 2 (Condition) x 2 (Time) x 4 (Sites) ANOVA revealed no significant main effect of condition or interactions. A step down analysis of the English group revealed a no main effect but a significant interaction of condition x time x site (F (3, 33) = 4.054; p = .032; partial eta squared = .269. A follow up analysis of sites will further investigate the interaction. An analysis of the left-posterior temporal site in a 2 (Condition) x 2 (Time) x group ANOVA revealed no main effect and a marginally significant interaction of condition x time x group (F (1, 22) = 4.275; p = .051; partial eta squared = .163). The same analysis examining the English group separately found no main effect but a significant interaction of condition x time (F (1, 11) = 6.623; p = .026; partial eta squared = .376). Post hoc comparisons revealed the condition differs for the later time interval (p < .01). Analysis of the left-anterior, right-anterior and right-posterior temporal sites each separately in a 2 (Condition) x 2 (Time) x group ANOVA revealed no significant main effect of condition or significant interactions involving condition.

Summary, CSD TN st and pt condition: In summary, a left-posterior TN was found in response to the *st* different target relative to the same (3-syllable) at 328 and 376 ms although the effect was greater at 328 ms. Language groups were not different.

A left-posterior TN was found in response to the *pt* different target relative to the same (3-syllable) at 328 ms and right-anterior TN at 328 and 376 ms. A language group effect was not found when examining each site separately.

Examining the Polish and English groups separately using data from all four temporal sites (AL, AR, PL, PR) at 328 and 376 ms revealed a minimal effect of language group because the *pt* conditions differed significantly for the 328 ms but not the 376 ms time interval for the Polish group. A similar interaction of condition x time was not found for the English group.

A language group effect was found for the 2-syllable *pt* contrast conditions. An interaction of condition, time, and group was revealed at the left posterior temporal site. Analysis of the English group data at the left posterior temporal site revealed the same and different targets to differ significantly at 376 ms.

CSD TN, comparison of st and pt difference waves 3-syllable: A comparison of the *st* and *pt* waveforms in a condition (st difference wave, pt difference wave) x time (280, 328, 376 ms) x site (PL, PR) ANOVA revealed no significant main effect and a significant condition x time interaction (F (2, 44) = 3.469; p = .04; partial eta squared = .136) showing the *pt* difference wave to be more negative than the st difference wave at 280 ms and more positive than the st difference wave at 328 and 376 ms. A significant interaction of condition x site (F (1, 22) = 6.24; p = .02; partial eta squared = .220) was also found as shown in Figure 2.21. Post hoc comparisons revealed that the *st* and *pt* difference waves did not differ significantly by time or site. The same analysis on the anterior temporal sites revealed no significant main effect of condition or significant interaction involving condition.

To further investigate the significant interactions at the posterior temporal sites a step down analysis of site was conducted. Analysis of the left and right posterior temporal sites separately did not find a significant main effect or interaction involving condition. Analysis of the Polish group separately in a condition (st difference wave, pt difference wave) x time (280, 328, 376 ms) x site (PL, PR) ANOVA revealed no main effect of condition or significant interaction involving condition. The same analysis on the English group data revealed no main effect but a significant interaction of condition x time (F (2, 22) = 4.74; p = .019; partial eta squared = .301) showing a more negative *pt* difference wave at 280 ms and a more negative *st* difference wave at 328 and 376 ms.

CSD T-Complex (Tb Component): *st* contrasts

Condition x site x group ANOVAs were carried out for the time bin 184 ms, where the Tb complex peaked as shown in Figures 2.22 and A.2.3.

CSD T-Complex (Tb Component) st contrast 3-syllable: A condition (2) x site (AL, AR, PL, PR) x group ANOVA revealed no main effect of condition or significant interaction involving condition. A condition x site interaction approached significance (p=.069). A follow-up analysis of the posterior sites in a condition (2) x site (2) x group ANOVA revealed no significant main effect and a significant interaction of condition x site (F (1, 22) = 4.959; p = .037; partial eta squared = .184). Post hoc comparisons of the condition x site interaction found the conditions to be significantly different only at the left posterior temporal site (p = .050) as shown in Figure 2.19. Follow up analysis of anterior sites did not find a significant main effect of condition or significant interactions.

CSD T-complex (Tb) st contrast 2-syllable: A condition (2) x site (AL, AR, PL, PR) x group ANOVA revealed no main effect of condition or significant interaction involving condition. An analysis of the anterior and posterior temporal sites follows to investigate hemisphere. Analysis of the posterior temporal sites revealed no significant main effect or interaction. Analysis of the anterior sites in a 2 (Condition) x 2 (AL, AR) x group ANOVA did not find a significant main effect of condition, but the interaction of condition x site x group approached significance (F (1, 22) = 3.976; p = .059; partial eta squared = .153) showing a more negative response to the different target by the English group from the left-anterior temporal site and a reverse response by the Polish group. A step down analysis of the left anterior and right anterior temporal sites separately revealed no significant main effect or interaction.

CSD T-Complex (Tb Component): *pt* contrasts

CSD T-Complex (Tb) pt contrast 3-syllable: A condition (2) x site (AL, AR, PL, PR) x group ANOVA revealed no main effect of condition or significant interaction involving condition. Anterior and posterior sites were examined separately to investigate hemisphere. No main effect of condition or interactions involving condition were found (2 (same, different) x 2 (PL, PR) x group; 2 (same, different) x 2 (AL, AR) x group.

CSD T-Complex (Tb) pt contrast 2-syllable: A condition (2) x site (AL, AR, PL, PR) x group ANOVA revealed no main effect of condition but a significant interaction of condition x group (F (1, 22) = 5.088; p = .034; partial eta squared = .188). Post hoc comparisons revealed that the Polish and English group response differs only for the different target (p < .05). A step-down analysis of data from each site follows to investigate the interaction. A 2 (Condition) x group ANOVA at the right-posterior temporal site revealed no main effect of condition but a significant interaction of condition x group (F (1, 22) = 6.639; p = .017; partial eta squared = .232), in which the Polish group showed a negative response to the different target relative to the same target and the English group showed the reverse pattern. Post hoc comparisons of the condition x group interaction revealed that the groups did not differ significantly for either the different target or the same target. A 2 (Condition) x group ANOVA at the left-anterior, the left-posterior and the right-anterior temporal sites separately revealed no significant main effect of condition or interaction involving condition.

Summary, *CSD T-Complex (Tb component) pt and st contrast:* In summary, CSD analysis of the Tb component revealed significant differences between the same and different pairs for *st* stimuli (3-syllable) from the left-posterior temporal site. A similar left-hemisphere effect was not evident for the *pt* stimuli. An effect of language group was revealed in response to the 2-syllable *pt* conditions, in which a more negative response to the different target (relative to the same) was evident for the Polish group and a reverse response for the English group from the right-posterior temporal site. Response to the 2-syllable *st* condition revealed an interaction of condition, group and site that approached significance and showed a more negative response to the different target by the English participants than the Polish participants from the left-anterior temporal site.

DISCUSSION

The purpose of the current experiment was to determine the role of the acoustic features of phonemes in speech perception. Acoustic features change with context so we questioned whether perception of phonemes (and phoneme segments) depends on the context of those phonemes in the native-language. To test this hypothesis, we examined two language groups (native-Polish and native-English) both of which have the consonant cluster /pt/ in the language, but only one group (Polish) experiences the cluster in the context examined in the current experiment (i.e., word onset). Specifically, we asked if native-English and native-Polish groups perceive the consonant cluster /pt/ in the same manner (behavioral task) and if the neural signal (ERP) reflects the perceptual response of the language groups. We compared perception and the neural response to the /pt/ consonant cluster with the response to the /st/ consonant cluster, which occurs in both languages in word onset. The /st/ stimuli served as an experimental control. Also, because the /st/ and /pt/ consonant clusters have different acoustic characteristics (e.g., duration) and because both clusters occur in word onset in Polish, we compared the influence of the differing acoustic characteristics on speech perception, irrespective of native-language exposure.

In the current experiment, the behavioral task revealed that Polish listeners were able to distinguish nonsense words having /pt/ and /pət/ in word onset, whereas English listeners were not. Having experienced word *endings* having /pt/ and /pət/ in the language was not sufficient for English listeners to accurately perceive the /pt/ in word onset nonsense words. This indicates that exposure to the phoneme and phoneme segments in the native-language results in an encoded neural signal that includes acoustic features of phonemes (or phoneme segments). Hence, lack of exposure to the /pt/ cluster in a specific context resulted in our English listeners being unable to perceive the word onset /pt/ cluster. The /st/ versus /sət/ contrast, which occurs in both languages in word onset, was perceived by both language groups with near perfect accuracy.

The ERP results revealed that late latency ERP responses reflected the behavior of the language groups. Specifically, an LPC peaking around 500 ms reflected the behavioral response of the language groups and a focal fronto-central positivity peaking around 400 ms, partially, reflected the language group response. Earlier latency ERP responses reflected the acoustic distinctions, both phonological and phonetic, distinctions, contrasting the stimuli. Furthermore, ERP results were more sensitive than behavioral results and provided fine-grained information about neural speech processing. ERP results revealed language group processing differences in response to the linguistic *st* contrast that were not evident from the behavioral results.

Native-Language Phonotactics

Phonotactic patterns of speech (e.g., phoneme sequences, stress patterns), including the distribution and frequency of the patterns, differ across languages (Halicki, 2010; Ewen & van der Hulst, 2001). From birth, infants are exposed to native-language phonotactic patterns within the speech signal and these patterns are neurally encoded early in development (Jusczyk et al., 1993; 1994; Saffran et al., 1996). Exposure to these patterns during development results in automatic processing of the native-language phoneme features (Strange & Shafer, 2008) suggesting that the neural encoding is instantiated at a pre-attentive level of processing. This hypothesis is supported by studies showing differences in native and non-native speech perception at the level of the mismatched negativity response. Automatic processing of native-language phoneme features should facilitate word recognition.

Native-language phonotactic patterns have been shown to affect speech production (Davidson and Stone, 2004). Dupoux et al., (1999) predicted that Japanese listeners' inability to accurately produce non-native consonant clusters would extend to the perception of those clusters. Dupoux and colleagues confirmed their prediction and demonstrated that Japanese listeners were unable to accurately perceive consonant clusters not present in the native-language, rather they perceived a vowel between the two consonants of the cluster (e.g., "ebzo" heard as "ebuzo"). In extending our knowledge pertaining to the influence of native-language phonotactic patterns on speech perception, we questioned whether a consonant cluster must occur in a particular context within the native-language to be perceived.

Native-Language Phonotactics, Behavioral Results

The phonotactic rules of English preclude /pt/ onsets and, as a result, English listeners were unable to perceive the word onset cluster, /pt/. Lack of exposure to the cluster and the sound contrast, /pt/ versus /pət/, resulted in the English listeners discriminating 2 and 3-syllable nonsense words at chance levels. In contrast, Polish listeners performed the task at well-above chance levels. As expected, perception of the *st* contrast, a contrast present in word onset in both languages, was not different for English and Polish listener groups. Berent et al., (2007), in a similar behavioral task, compared Russian and English listeners' perception of obstruent-obstruent word onset clusters, including /pt/, within nonsense words and compared them with contrast segments (e.g., /pət/) within nonsense words. Similar to our results, the Russian listeners were able to determine the number of syllables within the nonsense words, while English listeners, not having the onset clusters in their language, were not. We have demonstrated, along with Berent et al., and others, that the phonotactic patterns of the native-language influence speech perception (Dupoux, et al., 1999; Halle et al., 1998; Weber & Cutler, 2006). Furthermore, we have demonstrated that a phonemic segment (i.e., /pt/) must not only occur within the native-language but must occur within a particular context (e.g., word onset) to be perceived. The English listeners perceived the /pt/ pattern as a sequence within the native-language that was closest to the /pt/ onset cluster (i.e., /pət/ as in petunia), consistent with the research findings of Dupoux, et al., (1999) and Halle et al., (1998).

The native-English and native-Polish language groups heard the same stimuli and they were informed, through instruction and practice trials with feedback, that both 2 and 3-syllable nonsense words would be presented. Nevertheless, the English listeners were unable to distinguish the 2 and 3-syllable *pt* nonsense words. Listeners contrasting non-native phonemes are said to use a "phonetic mode of processing that requires attentional resources" (Strange & Shafer, 2008, p 174). Our syllable identification task might have been too challenging for the English listeners to access the phonetic code but it is possible that under different task conditions non-native participants would be able to access this information. Also, Strange and Shafer (2008) argue that native-phoneme contrasts are perceived automatically even in poor listening conditions. The native-Polish listeners reported that the speaker of our experimental stimuli, while having good Polish pronunciation was a dominant English speaker. Even in this more challenging situation, the Polish listeners correctly identified the 2 and 3-syllable *pt* nonsense words (linguistically relevant in Polish), while our English listeners did not.

Acoustic features of stop consonants (i.e., /p/, /t/, /k/) in word onset differ in Polish and English. Polish speakers release stop consonants after closure without aspiration, whereas, English speakers release and aspirate these stop consonants after closure in /pət/ as in *petunia* and in /sət/ as in *sateen* (Ladefoged, 2001, p 43). The /t/ phoneme in /st/ is released and not aspirated in both Polish and English phonology (personal communication with Katarzyna Dziubalska-Kolaczyk). The difference in the acoustic features of the stop consonants in English and Polish did not aid our English listeners. Our analysis of response error by the Polish participants revealed that the presence or absence of a vowel was a factor in perception of the *pt* contrast. Twenty 3-syllable *pt* tokens did not have a perceptible vowel between the /p/ and /t/ and were identified, incorrectly, as having 2-syllables by the subgroup of Polish participants examined. In contrast, all 3-syllable nonsense words (within 243 trials) correctly identified as having 3-syllables by the Polish listeners, had a perceptible vowel. We do not claim to have established which acoustic features are used by the native-Polish listeners to recognize the 2 and 3-syllable *pt* nonsense words, but have established that the English group, presented with the same stimuli as the Polish group, were unable to differentiate the 2 and 3-syllable *pt* words.

The English language input contains the cluster /pt/ (e.g., He sle*pt in a* bed), but because it is an illegal form in word onset, it is never heard without a preceding vowel. Consider that 3-syllable words beginning with /pət/ may approximate 2-syllable words within rapid speech. For example, the schwa between the /p/ and /t/ in the word *potato* might be reduced in rapid speech so that the initial segment approaches the consonant cluster /pt/. Consider the sentence, “It is a big potato.” In rapid speech, the initial segment could potentially be reduced to approach the onset cluster /pt/. However, even in the case of this potential English input, the stop consonant, /p/, would retain the feature aspiration. Deletion of the schwa in an onset pre-stressed syllable (e.g., 3-syllable *pt* word forms, “petunia”) is rare in English (Peterson, LoCasto & Connine, 2002). However, Peterson & colleagues found that when a competitor exists in the language for the reduced form (i.e., “sport” is a competitor for the reduced form of “support”) the schwa is less likely to be deleted in high frequency words. Because /pt/ is an illegal form in English, there are no competitors for a reduced form of the 3-syllable *pt* word (e.g., “petunia”). It is possible, however, that Polish speakers delete the schwa in 3-syllable *pt* words less frequently than English speakers because competitors of the reduced forms are possible. This fact provides an explanation for the speaker of our experimental stimuli, who is a dominant English speaker, producing a small set of reduced variants for the 3-syllable *pt* words that were incorrectly perceived by the native-Polish participants.

Nevertheless, the current results indicate that if reduced syllables, such as described, occur in English, they do not aid the English listeners in recognizing the *pt* contrast. Either the different acoustic features in the English and Polish stop consonant /p/ or limited experience with the reduced segments fail to aid the English listeners. Future research may provide answers concerning the specific acoustic features that permitted the Polish group, but not the English group to recognize the *pt* contrast.

Native-Language Phonotactics ERP Results

Speech perception abilities of the native-language groups, revealed through behavioral measures, were reflected in a Late Positive Component (LPC) and, partially, in a focal fronto-central positivity at 400 ms. Detection of the *st* contrast, linguistically relevant in Polish and English, was reflected in the LPC component by both language groups. In contrast, the LPC response to the *pt* sound contrast, which is linguistically relevant in word onset in the Polish language only, was different for the language groups The English listeners' LPC response to the *st* difference wave was found to be different from their response to the *pt* difference wave, but the Polish listeners' LPC response was not.

Late processing LPC

The LPC response in the current experiment was consistent with a P3b type response thought to reflect evaluation of a stimulus for an overt response (Nobre and McCarthy, 1994; O'Rourke & Holcomb, 2002) and has been reported, previously, in response to a phonological contrast (Schwartz et al., in preparation; Praamstra et al., 1993). Using the phonological priming design (same and different word pairs having a short ISI) allowed us to determine if the contrast in the different pair elicited an LPC response, substantiating the behavioral response. Based on previous research findings, we argue that the LPC reflected the participants' conscious evaluation of the target word stimuli for responding on the syllable identification task. The GFP that differed significantly for the *st* contrast conditions around 500 ms corresponded to a posterior-parietal positivity in the CSD maps supporting the existence of a late positive component. Native-language patterns of speech perception were reflected within the LPC and, to our knowledge, has not previously been used as a means to compare perception of specific onset consonant clusters in two language groups.

A study by Dehaene-Lambertz, Dupoux & Gout (2000) investigated French and Japanese listeners' ERP response to a vowel contrast within nonsense word stimuli (i.e., 6 phonemic forms including "igmo" versus "igumo"). An oddball task design (600 ISI between words) was used and participants' responded with a same/different decision to the final word in a train of 5 words. In agreement with our results, native-language phonotactic patterns were reflected in the LPC response and in an earlier negative response around 400 ms. Our study did not show this negative response possibly because of design differences. Rather a frontal positivity was observed around 400 ms (P400).

P400

A focal fronto-central positivity, peaking around 400 ms was found to, partially, reflect the influence of the native language. Both English and Polish language groups demonstrated neural discrimination of the *st* and *pt* contrasts (i.e., /st/ from /sət/ and /pt/ from /pət/ in word onset) as indexed by a large P400 for different compared to same targets. The P400 latency in response to the *pt* contrast, however, was significantly different for the language groups at the fronto-central site with an earlier response by the Polish group. Interestingly, the response to the *st* contrast (3-syllable) examining central, right and left frontal sites, found the conditions to be significantly different, only at the fronto-central site. Our findings demonstrate that the amplitude of the P400 component indexes acoustic aspects of the contrast because, both, the *st* and *pt* contrasts are recognized by the English group. In contrast, latency of the P400 reflects native-language experience because latency for neural processing is significantly earlier for the Polish group in response to the *pt* contrast (*pt* difference wave; 3-syllable).

The study by Dehaene-Lambertz and colleagues (2000) found a negative component to a vowel contrast for French participants, but not Japanese participants, between 384 and 488 ms from fronto-central sites consistent with the current results. The results of the two studies differed, however, because we found a native-language effect of latency but not amplitude and polarity of the waveforms was not the same. The different findings across studies can be explained by experimental design differences. Furthermore, CSD transformations can reveal components that are masked by deep source activity.

A recent study having a word-fragment priming design and a lexical decision task found a phonological N400-type response, peaking at posterior central sites around 350 ms, in response to their different word condition (e.g., hun-dragon) but not to a change of one onset phoneme (Friedrich et al., 2009). In contrast, the less distinct variation consisting of a change of one onset phoneme was recognized at 350 ms from the left lateral region. In the current study, we found a fronto-central positivity, as an effect of priming, in response to the contrast within the different pairs. The experimental design used by Friedrich et al., differed from ours because they used a lexical decision task with a word fragment as the prime as well as different contrast types. Below, we will discuss further the failure to observe an N400 type effect.

ERP Timecourse of the Native-Language Effect

An effect of native-language phonotactic patterns was evident within the ERP waveform from ~ 400 ms demonstrated by group latency differences (P400 component) from the fronto-central electrode site in response to the *pt* contrast. A later native-language effect was found reflected in the LPC peaking around 500 ms. Recall that the vowel contrasting our different word pairs was the second phoneme within the first syllable. The earliest effect of native-language phonotactics within the Dehaene-Lambertz et al., (2000) experiment also occurred ~ 400 ms. from fronto-central electrode sites in response to a medial vowel contrast within nonsense word stimuli (e.g., “igmo” versus “igumo”).

A minimal effect of language was also revealed within the TN and Tb components. Each lateral temporal site tested separately did not find an interaction of condition and language group for the 3-syllable *pt* contrast conditions. However, Polish participants’ 3-syllable *pt* contrast data from all four lateral temporal sites at 328 and 376 ms revealed a greater effect of condition at 328 ms. A similar effect was not found in the English participant data. Also, 2-syllable *pt* conditions differed at 376 ms (left posterior temporal site) for the English participant data only. The earlier time bin (328 ms) showed a greater effect of condition for the 3-syllable *pt* conditions suggesting that the language group effect for the 2-syllable condition may not be reliable or may reflect increased variance in the English group response. A minimal effect of language was also found at 184 ms in response to the *pt* contrast (2-syllable) from the right-posterior temporal site. A native-language effect was not found within the 3-syllable condition, suggesting that the finding is not robust because significant results that appear in the 2-syllable conditions are generally found within the 3-syllable conditions.

ERP Fine-Grained Language Effect

We predicted both behavioral and late latency ERP responses to be different for our English and Polish participants in response to the *pt* contrast but to be the same in response to the *st* contrast. The behavioral results were consistent with this prediction, whereas, the ERP responses only partially confirmed this prediction.

Late latency ERP responses showed the expected pattern to the *pt* contrast. The English and Polish listeners performed differently on the behavioral task to the *pt* contrast and this difference was evident in the LPC component of the ERP waveform. The behavioral response to the *st* contrast was the same for both language groups (near perfect accuracy). The LPC reflected detection of the *st* contrast for both language groups, however, the response to the *st* contrast was larger for the English listener group than the Polish listener group and topographical group differences for this component were found. Both language groups showed a greater LPC to the different target (relative to the same) from the midline-posterior parietal sites but only the English group showed a greater LPC to the different target from the right-posterior parietal sites. Also, the Polish participants showed relatively large variability compared to the English participants in response to the *st* conditions whereas the English participants showed relatively large variability compared to the Polish participants to the *pt* conditions.

The current research has demonstrated that perception of native and non-native contrasts can be reflected in the LPC. As previously stated, the LPC reflects multiple cognitive processes (Steinschneider & Dunn, 2002; Picton, 1992), therefore, processes in addition to phonological detection may be reflected within the LPC. For instance the *st* contrast, present in both languages, may have been more challenging for our Polish group because our native-Polish speaker of the experimental stimuli was a dominant English speaker. While not evident in the behavioral results because both groups demonstrated near perfect accuracy, the *st* stimuli spoken by a Polish-English bilingual who is dominant in English may have required greater cognitive resources by the Polish participants.

Johnson (1986) discusses three aspects of stimulus meaning that can contribute in an additive fashion to P3 amplitude: (1) task complexity, (2) stimulus complexity and (3) stimulus value providing an alternative explanation for the greater LPC to the *st* contrast for the English listeners. It is possible that a phoneme sequence such as word onset /st/, has a high stimulus value in English because frequently used and early-learned words begin with that phoneme segment (e.g., stop). Furthermore, a P3 vertex response has been shown to reflect greater confidence in detection of an auditory signal (Squires et al, 1974) suggesting ease of processing by the English listeners for high frequency phoneme sequences occurring within a language as a possible explanation.

Whether the LPC, larger to the *st* contrast for English listeners, reflects greater frequency of usage of /st/ and /sɚt/ words in the English language, greater stimulus value or ease of processing awaits further research. Subtle differences in the distribution of sounds within words have been found to affect spoken word recognition (Vitavitch, 2007) and, therefore, distributional differences of phoneme sequences within languages may affect neural speech processing. To our knowledge, language group differences within the ERP response to a word onset *st* contrast have not been previously demonstrated and, therefore, the results until replicated, along with speculation concerning the results, must be considered with caution.

Acoustic Features

We have demonstrated through behavioral results and late latency ERP responses that acoustic features of phonemes influence perception. The language exposure of the native-Polish listeners during development includes the acoustic features of the word onset cluster /pt/ and the contrast segment /pət/, shaping perception of the sound sequences. The English native-language experience, in contrast, does not contain the onset cluster /pt/ with its contextual features and, therefore, does not become part of the neural code in perceptual speech processing. Having the cluster /pt/ and the contrast segment in the English language in another context did not aid in perception of the onset cluster indicating that the contextual features of phonemes are intrinsic to perceptual speech processing. We have proposed the acoustic features of phonemes as the basic unit of words.

A model having the acoustic feature as the basic unit of words is consistent with neurophysiological studies of speech processing, theories of lexical activation and developmental research findings. Speech is processed by acoustic features in primary auditory cortex (Steinschneider et al., 1995; Steinschneider et al., 2005). The features are then, in some way, integrated into word-like units for meaning to be attached. Current research suggests synchronous neural activity as a potential means of integrating neural responses (Eggermont, 2005). Synchronous (millisecond) neural firing in response to acoustic features of sounds could integrate the features of the phoneme sequence (word) as they change over time. Zhang, Kuhl, Imada, Kotani & Tohkura (2005) argued that acoustic features of the native-language syllable /la/ were integrated for American, but not Japanese participants. The researchers conducted a MEG (magnetoencephalography) study in which American and Japanese participants were presented with the endpoints of

/ra-la/ and /ba-wa/ continuum using an oddball experimental design. The /ra-la/ stimuli contained two portions of the third formant each having "well-defined onsets" (p 710); The initial portion of F3 contained 155 ms of steady state sound and the second portion, which was critical for the /ra-la/ distinction, consisted of a 100 ms transition. Japanese listeners failed to integrate the two portions of the F3, as evidenced by a double $N1_m$ component within the right and left hemispheres in response to /la/. In contrast, the American listeners integrated the two onsets of /la/, as evidenced by only one clear N1m within the left hemisphere. The authors argued that native-language experience alters the temporal integration time-window because Japanese listeners did not integrate the late F3 transitions with the earlier acoustic features (including F3 steady state sound) of the /la/ stimulus. An earlier MMF (mismatch field, the magnetic counterpart of the mismatch negativity) for the Japanese listeners in response to the /ra-la/ contrast supported this view.

The proposal that acoustic features serve as a basic unit of the word is consistent with current theories of lexical activation, which have recognized that acoustic features of phonemes affect word recognition (McQueen et al., 2003 Andruski et al., 1994). Also, languages have a limited number of sounds and a large variety of words that can be accessed from sound combinations (McQueen et al., 2003). Acoustic features as the underlying basic unit of words allows for a feasible word recognition process because they constrain the possibilities in word recognition.

Furthermore, developmental research findings are consistent with acoustic features as an intrinsic aspect of perceptual speech processing. The sound patterns and frequency of sound patterns of the native-language are coded in the neural signal before word meanings are learned indicating that formation of the native-language perceptual patterns precede meaningful linguistic (phonological) contrasts (Jusczyk et al., 1993; 1994). Eight-month old infants' detection of 3-syllable nonsense words from a stream of syllables after 2-minutes of exposure (Saffran et al., 1996) can be explained because the acoustic features change with every phonemic context so there is only one acoustic combination for a word. Fast mapping of acoustic features of the phoneme sequence on auditory neural networks could explain the findings of Saffran and colleagues. Babies as young as 13-months of age comprehend novel words after hearing the words only nine times, a process termed "fast mapping" (Woodward, Markman & Fitzsimmons, 1994). It has been argued that infants use statistical transitional probabilities to learn the order of syllables, an inferred cognitive skill, as a means for detecting novel syllable sequences (Saffran et al.,1996; Kuhl, 2000). A cognitive skill is not necessary because the unique acoustic sequences of words can be mapped as the auditory neural code. Early development of the auditory cortical regions relative to the prefrontal cortical regions, which are involved in higher-level cognitive function, is consistent with this theory (Huttenlocher & Dabholkar, 1997).

Integration of acoustic features of phonemes into word-units may be the sound-based representation that not only captures phonological information as stated by Praamstra & Stegemann (1993), but is the phonological information in that it is the native-language input and through frequency of the input during development becomes the speech signal that is perceived rapidly and automatically.

ERP Components

ERP data is interpreted within the context of the behavioral task and the phonological priming design. Recall that the ERP epochs are time-locked to the onset of the target word of the same and different pairs and the target words are unchanged in the same and different condition. When the target word is repeated by using a different token within the same pairs and an ISI of 250 ms, neurons are in a refractory state and neural firing in response to the sounds and features of the same word are reduced (Sussman, Steinschneider, Gumenyuk, Grushko, & Lawson, 2008). In the different conditions, new neural firing occurs in response to sounds and features of the target word that are different. A stronger ERP response was found for the 3-syllable conditions (i.e., a 3-syllable target word was preceded by a 2-syllable word) relative to the 2-syllable conditions (i.e., a 2-syllable target word was preceded by a 3 syllable word). Because there is a short interval between the words in a pair, a 3-syllable word following a 2-syllable word in the different pair results in greater neural firing, at least in part, because there are a greater number of sounds and features in the longer word. A 2-syllable word following the 3-syllable word, on the other hand, finds a shorter word as the second word with fewer sounds and features. Language group and condition differences in the ERP response may have been less robust in these conditions because less recovery from refractoriness resulted in a smaller signal.

P400: A response to the physical contrast in the *pt* and *st* stimuli was evident in both language groups in the form of a focal fronto-central positivity at 400 ms (P400). The earlier fronto-central sensory-obligatory components (P1N1P2) within the target word data did not reflect the physical contrast within the different pairs. It is possible that these obligatory components did not index differences between the contrasts because the short ISI between words within the pair resulted in a small P1N1P2 response and, thus, differences were masked by noise. The longer latency P400 response may reflect detection of a difference within the word unit. A fronto-central response with an inversion at the mastoids is consistent with a source in auditory cortex (Ponton et al., 2002).

N400: The N400 response, predicted to reflect the linguistically-relevant sound contrasts within words, was not evident in our ERP data. It is possible that the N400 was masked by the large LPC component (Holcomb et al., 2002). Alternately, the N400 component may have been absent because participants were aware that all stimuli were nonsense words and, thus, did not attempt to access lexical representations. Praamstra and Stegeman, 1993 found an N400 to be absent when participants could make a lexical decision without listening to the whole word. The nonsense word onsets in Praamstra et al., were illegal consonant clusters, so participants were able to determine that the stimulus was not a real word from the onset. They found an N400 component in response to the same experimental stimuli when a rhyming task was used. Studies using similar experimental designs to ours that found an N400 response presented both real and nonsense word stimuli (Praamstra & Stegerman, 1993; Shafer, Schwartz & Kessler, 2002; Bonte & Blomert, 2004, Friedrich et al., 2009; Schwartz et al., in preparation).

Tb component of the T-Complex and TN: The posterior temporal negativity found at 184 ms is consistent with the Tb component of the T-Complex (Ponton et al., 2002). The Tb component consisted of a left-hemisphere response for the *st* sound contrast that was not evident for the *pt* sound contrast. A left-hemisphere effect for the *st* contrast and a bilateral effect for the *pt* contrast for the later TN response were also found. Both *st* and *pt* contrasts are phonological contrasts in the Polish language. A left-hemisphere effect for one sound contrast and not the other suggests that acoustic aspects of the stimuli are processed differently.

Psychoacoustic Salience

Acoustic features of phonemes affect native-language speech perception (Burnham, 1986; Berent et al., 2007; Strange & Shafer, 2009; Jusczyk et al., 1993). Burnham proposed that "psychoacoustic saliency" of phoneme contrasts influences the development of native-language perception, influences adults' ability to distinguish non-native contrasts after training and underlies whether sound contrasts are found to be common or rare in the world's languages. In the current experiment, the ERP response from lateral temporal sites differed in response to the *st* and *pt* contrasts, which could not be explained by native-language experience because the effect occurred for both English and Polish language groups.

The /pt/ cluster consists of two voiceless stop consonants of short, rapidly changing, acoustic features and /st/ consists of a fricative and stop consonant, which have a different manner of articulation. The phoneme /s/ is a long duration consonant having high-energy frication noise. In contrast, the phoneme /p/ is a consonant having low energy and rapidly changing acoustic features (see Bordon et al., 2003).

The clusters (i.e., /pt/ and /st/), the syllables (i.e., /pət/ and /sət/) and the schwa between the segments were found to have different mean durations, although the difference in mean duration for the schwa was only 9 ms. Examination of the grand mean waveforms in response to the *st* and *pt* contrasts from the left-posterior temporal site in Figure 2.22 reveals the peak of the TN waveform to differ for the sound contrasts by 70 ms, suggesting an ERP response involving more than the vowel contrast. The ERP response, therefore, within the context of the phonological priming task may not only be a response to the contrast segment (i.e., the schwa) but a response to the integrated acoustic information. Acoustic elements may be integrated when occurring within the temporal integration window of ~ 200 ms and, therefore, the contrasting acoustic features of the /s/ and /p/ phonemes (duration and presence or absence of frication noise) may be integrated within the temporal integration window affecting the temporal ERP components (Winkler, Czigler, Jaramillo, Paavilainen & Naatanen, 1998).

As discussed earlier, Friedrich et al., (2009) did not find an effect of a contrast in one onset phoneme (i.e., variation condition) for the T-Complex response using a priming task. The authors argued that the T-Complex did not register linguistic information. In contrast, we found an effect of condition (a vowel contrast) within the T-Complex response for the *st* conditions but not the *pt* conditions, both linguistic contrasts in Polish. The results of the two studies may differ because Friedrich et al., examined different phoneme combinations (i.e., onset stops and nasal consonants examined within the same experimental conditions) and we examined specific phonemic combinations. Also, tasks differed in that Friedrich et al., used a lexical decision task and our participants identified syllables within nonsense words.

Research findings pertaining to a left hemisphere effect for phonetic speech processing have been inconclusive (Zhang et al., 2005; Rinno, et al., 1999; Alho et al., 1998). Similar to our results, Zhang and colleagues found American participants to demonstrate left-hemisphere processing in response to a /ra-la/ contrast but not in response to a /ba-wa/ contrast. Interestingly, the /ra-la/ continuum has continuent consonants of long duration (relative to stop consonants) and the /ba-wa/ continuum contains a stop and a continuent consonant. It is possible that the nature of the distinction, spectral versus temporal, influences involvement of the left hemisphere.

The /st/ consonant cluster is assumed to be more perceptually salient than the /pt/ consonant cluster. The /st/ consonant cluster is common in the world's languages and is an exception to principles of sonority and to principles for word onset consonant cluster use in English. For example, /st/ is an exception to illegal word onset obstruent-obstruent sequences (e.g., **ft, *pt, st*) and to illegal word onset homorganic sequences (e.g., **pw, st*) in English (Gierut, 1999). In contrast, the cluster /pt/ is rare in the worlds' languages and is rare in the Polish language. The current ERP results provide neurophysiological support for the view that the /st/ cluster has greater perceptual salience than the /pt/ cluster. Furthermore, our results confirm that specific sound contrasts must be examined in order to tease apart the effects of the native language and acoustic factors on speech perception, as suggested by Strange & Shafer (2009) and Jusczyk et al., (1993).

Conclusion

Our data support the claim that acoustic features of phonemes (i.e., contextual features) are intrinsic to perceptual speech processing. Exposure to phonotactic patterns within the native language during development results in an encoded neural signal that reflects the acoustic features of phonemes found in a listener's native-language. The results suggest that acoustic features are necessary for word recognition. Our data are consistent with neurophysiological results of speech processing, lexical activation research and infant models of developmental speech perception.

Affects of native-language experience were observed within the ERP signal late in processing beginning around 400 ms post-word onset. Earlier latency ERP components revealed that the acoustic distinctions within the experimental stimuli are registered in the brains of both language groups, in the same manner. This finding will be expanded upon in Chapter 3 in which we report the prime word data results.

Also, results of the current study revealed that neural correlates of speech processing differed for the *st* and *pt* stimuli, for both the English and Polish listeners, irrespective of native-language. These results provide neurophysiological support for the view that the *st* contrast is more salient than the *pt* contrast, consistent with the notion that /pt/ is more marked (rare) than /st/ in the world's languages. Furthermore, a left-hemisphere effect for the *st* sound contrast, but not the *pt* sound contrast, suggests that acoustic features of sounds should be considered to understand the role of the left hemisphere in the processing of speech.

Table 2.1
Eight experimental conditions: 4 *pt* and 4 *st* experimental conditions

CONDITION TYPES	***pt***	***st***
SAME 2	/ptukɑ-ptukɑ/	/stɛsɑ-stɛsɑ/
SAME 3	/pətimɑ-pətimɑ/	/sətIlɑ-sətılɑ/
DIFFERENT 2	/pətukɑ-ptukɑ/	/sətɛsɑ-stɛsɑ/
DIFFERENT 3	/ptimɑ-pətimɑ/	/stılɑ-sətılɑ/

Table 2.2 Mean percentage correct scores, median percentage correct scores and standard deviation scores are shown for the syllable identification task

CONDITIONS	ST CONDITIONS		PT CONDITIONS	
	Polish	English	Polish	English
SAME 2				
Mean	**98**	**97**	**90**	**43**
Median	**99**	**99**	**97**	**49**
SD	**1.8**	**4.7**	**18.4**	**33**
DIFFERENT 2				
Mean	**99**	**98**	**94**	**42**
Median	**99**	**99**	**98**	**50**
SD	**.9**	**3.2**	**8.4**	**30.1**
SAME 3				
Mean	**99**	**98**	**81**	**85**
Median	**100**	**99**	**82**	**90**
SD	**2.4**	**2.2**	**7.9**	**20.4**
DIFFERENT 3				
Mean	**99**	**99**	**84**	**82**
Median	**100**	**100**	**83**	**88**
SD	**1.1**	**1.8**	**5.6**	**20.8**

Table 2.3 depicts individual participants' percentage correct scores on the syllable identification task for the *st* conditions. The word "same" or "diff" in the condition label refers to same or different pairs and the number following the word refers to the number of syllables within the target nonsense word. Chance level performance equals 50%.

ST CONDITIONS: INDIVIDUAL PERCENTAGE CORRECT SCORES
SYLLABLE IDENTIFICATION TASK

ENGLISH	*Same2*	*Diff2*	*Same3*	*Diff3*	**POLISH**	*Same2*	*Diff2*	*Same3*	*Diff3*
1	100	100	99	100	**1**	99	98	100	100
2	97	99	97	99	**2**	99	99	99	98
3	97	94	100	100	**3**	100	100	100	100
4	97	100	100	99	**4**	94	97	94	97
5	99	100	98	100	**5**	99	99	100	99
6	89	100	99	99	**6**	98	100	100	100
7	90	98	99	100	**7**	97	100	99	100
8	100	99	100	100	**8**	99	99	100	100
9	100	100	92	99	**9**	100	100	100	100
10	87	89	96	95	**10**	96	99	99	98
11	100	98	100	95	**11**	99	100	93	99
12	100	99	98	100	**12**	99	99	99	100
13	100	99	99	100					

Table 2.4 depicts individual participants' percentage correct scores on the syllable identification task for the *pt* conditions. The word "same" or "diff" in the condition label refers to same or different pairs and the number following the word refers to the number of syllables within the target nonsense word. Chance level performance equals 50%.

PT CONDITIONS: INDIVIDUAL PERCENTAGE CORRECT SCORES SYLLABLE IDENTIFICATION TASK

ENGLISH	*Same2*	*Diff2*	*Same3*	*Diff3*	**POLISH**	*Same2*	*Diff2*	*Same3*	*Diff3*
1	73	50	79	72	**1**	33	75	92	99
2	51	53	85	79	**2**	97	99	75	82
3	0	0	100	99	**3**	94	96	88	84
4	49	60	94	83	**4**	89	95	81	85
5	85	70	35	32	**5**	92	98	89	86
6	1	1	100	100	**6**	99	98	89	90
7	54	53	90	90	**7**	97	98	82	83
8	76	82	77	85	**8**	96	99	69	82
9	90	92	51	49	**9**	97	98	80	83
10	18	15	100	94	**10**	87	78	87	80
11	3	20	100	100	**11**	100	98	71	80
12	11	16	99	99	**12**	98	98	73	78
13	49	39	89	88					

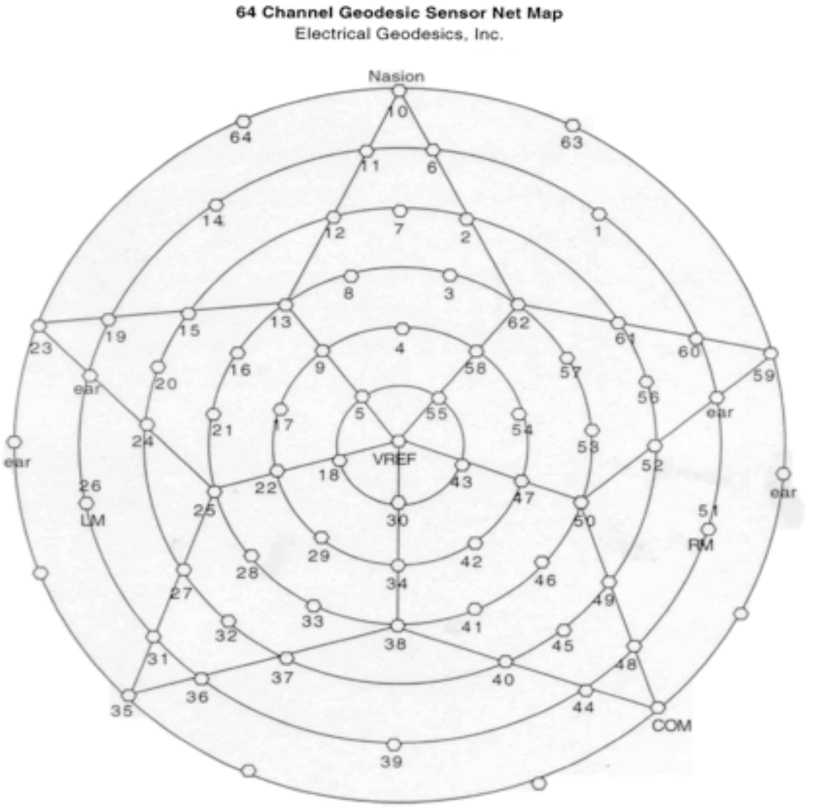


Figure 2.1 shows a map of Electrical Geodesic Inc. electrode sites

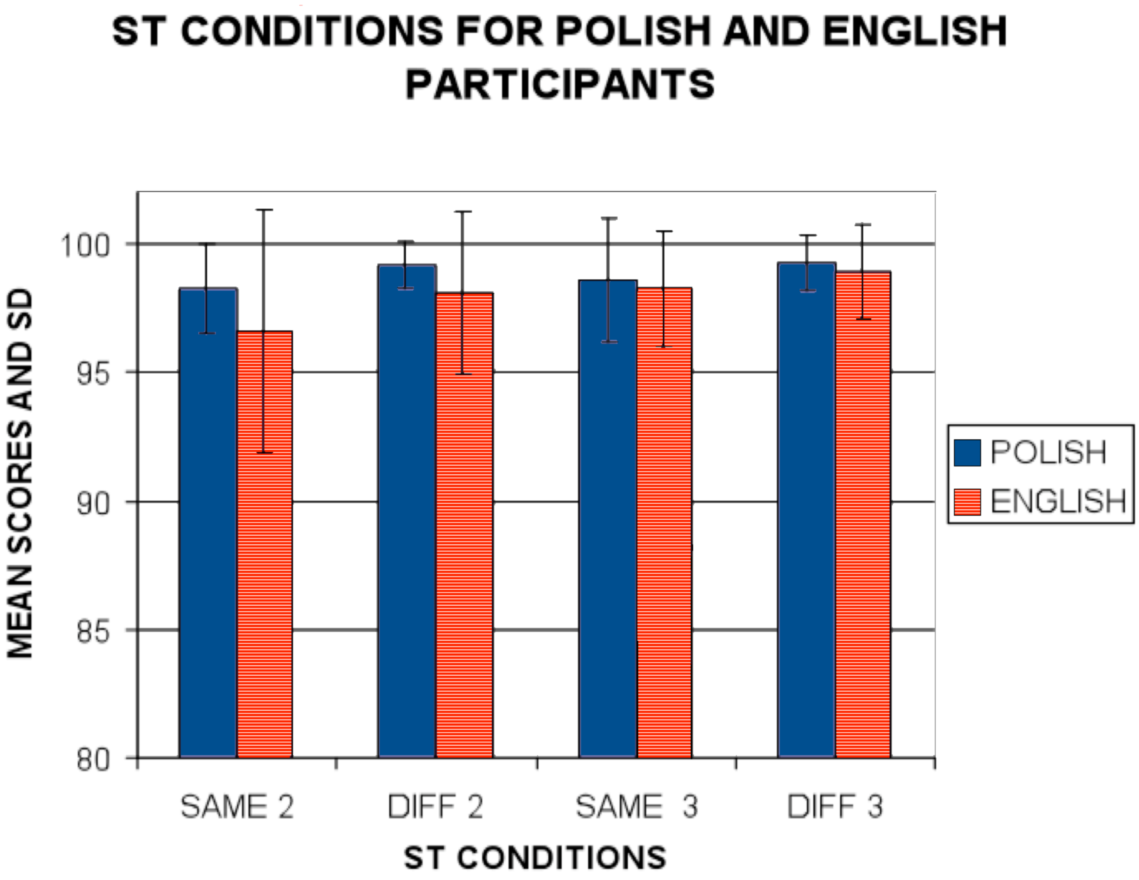


Figure 2.2 depicts the mean scores and standard deviation for the *st* conditions. Note that the scale is different for the *pt* and *st* charts

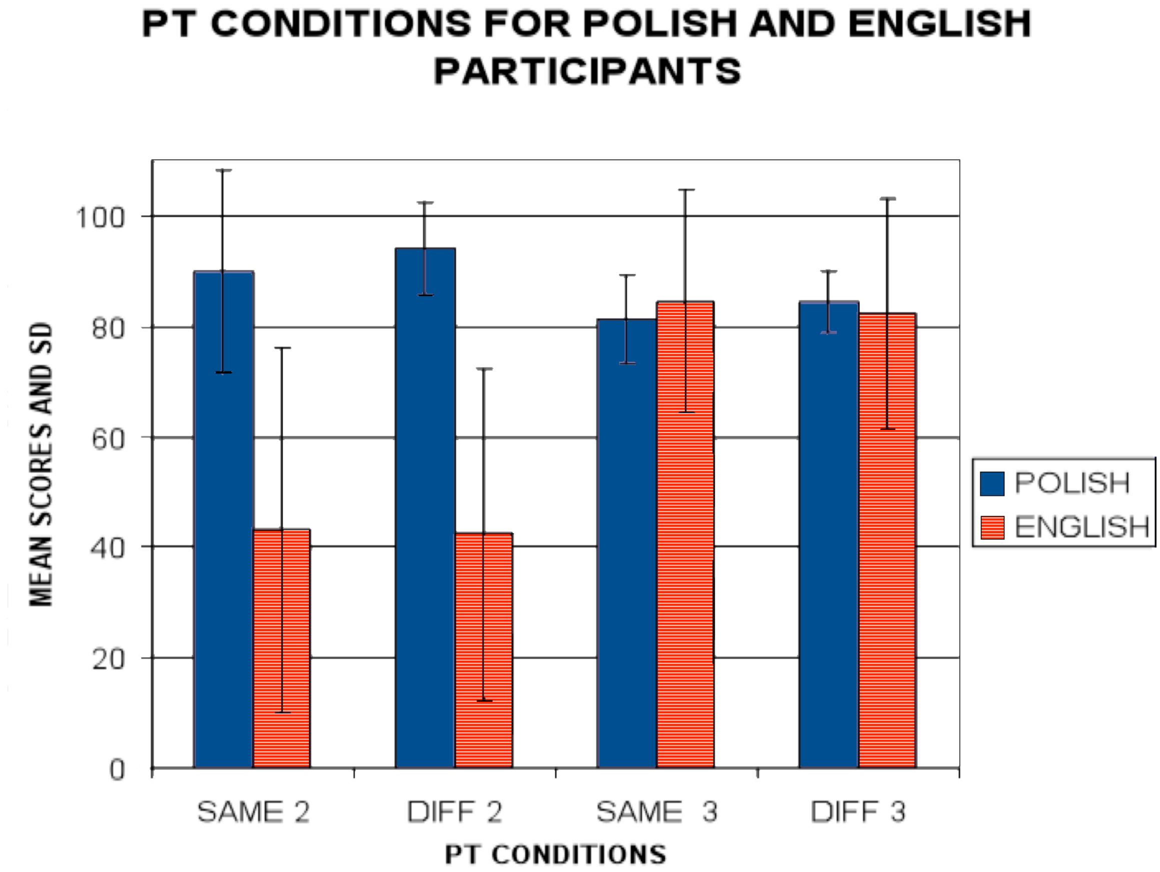
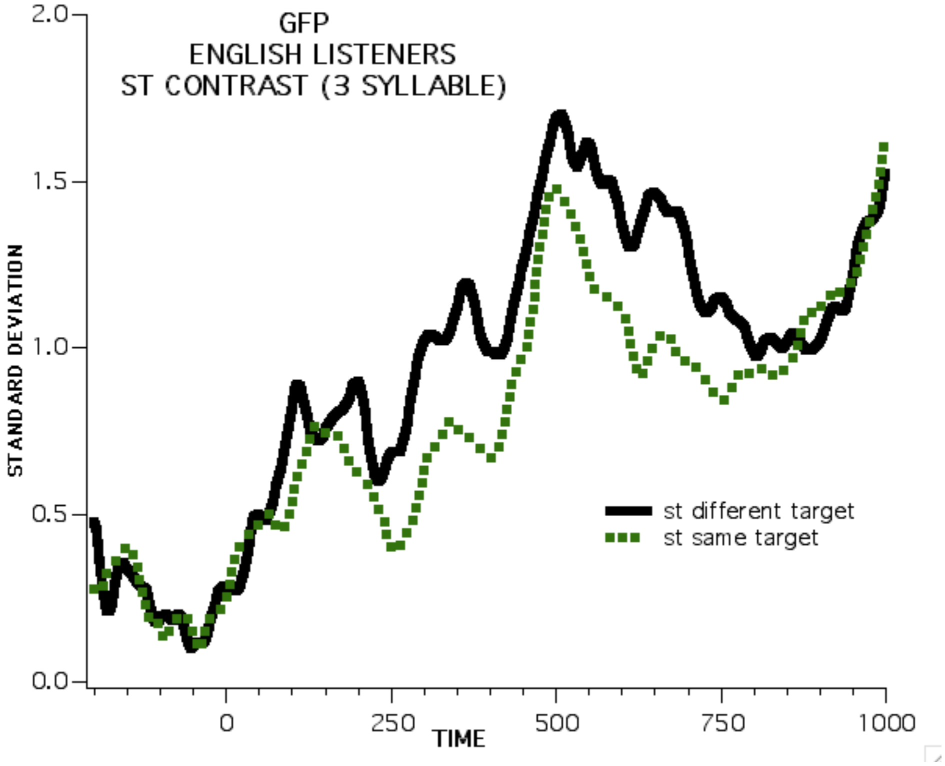


Figure 2.3 depicts the mean scores and standard deviation for the *pt* conditions. Note that the scale is different for the *pt* and *st* charts



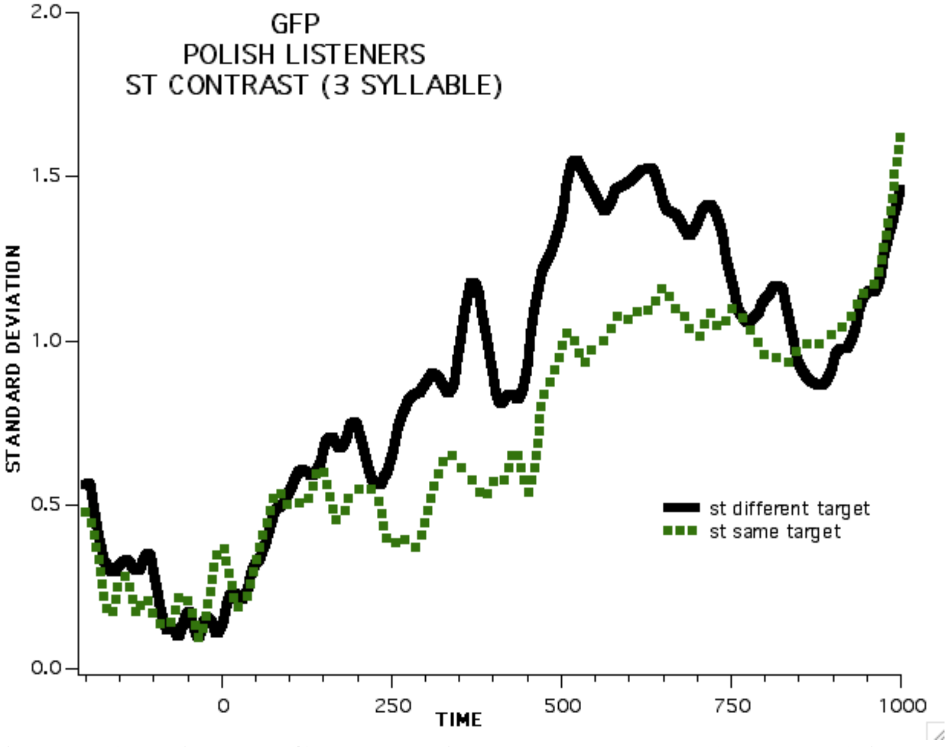
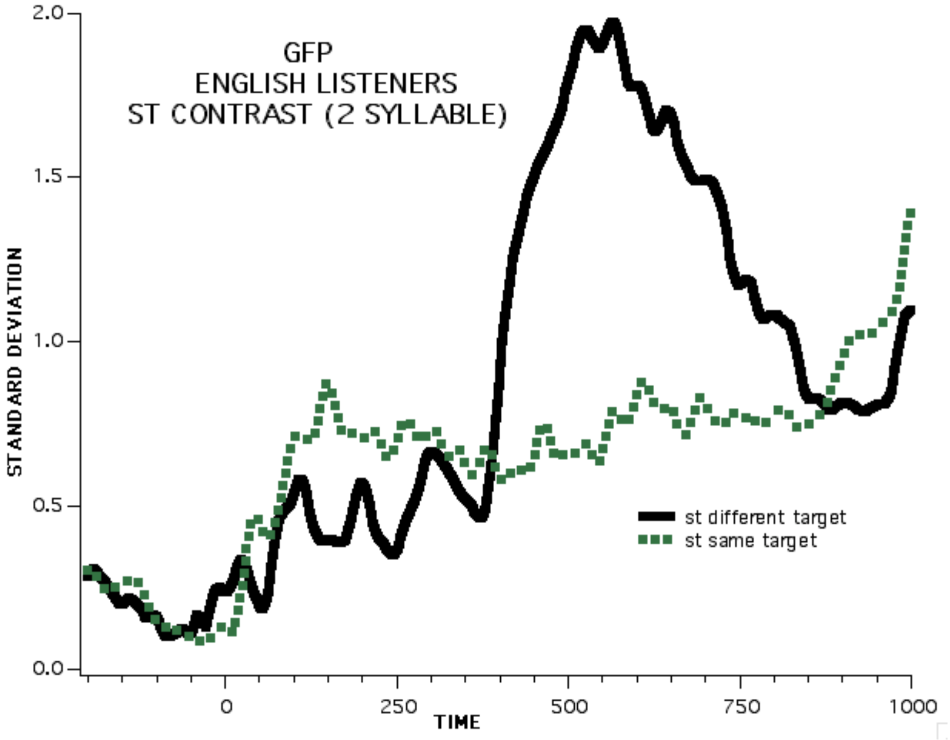


Figure 2.4 depicts the GFP analysis to the *st* contrast (3-syllable) in English (top) and Polish (bottom) participants



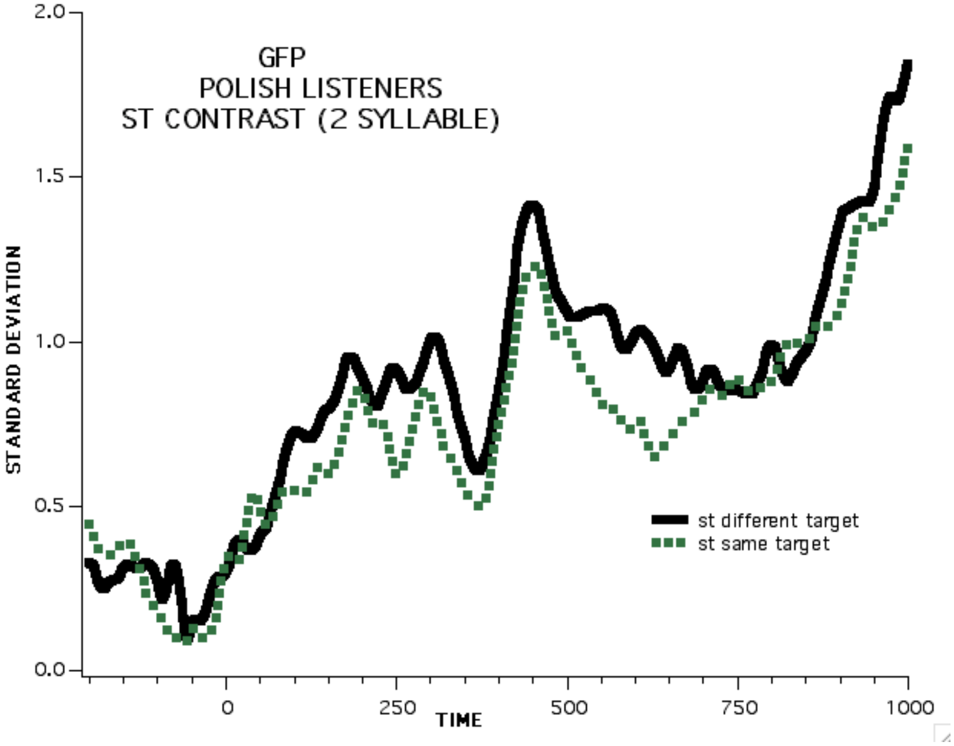


Figure 2.5 depicts the GFP analysis to the *st* contrast (2-syllable) in English (top) and Polish (bottom) participants

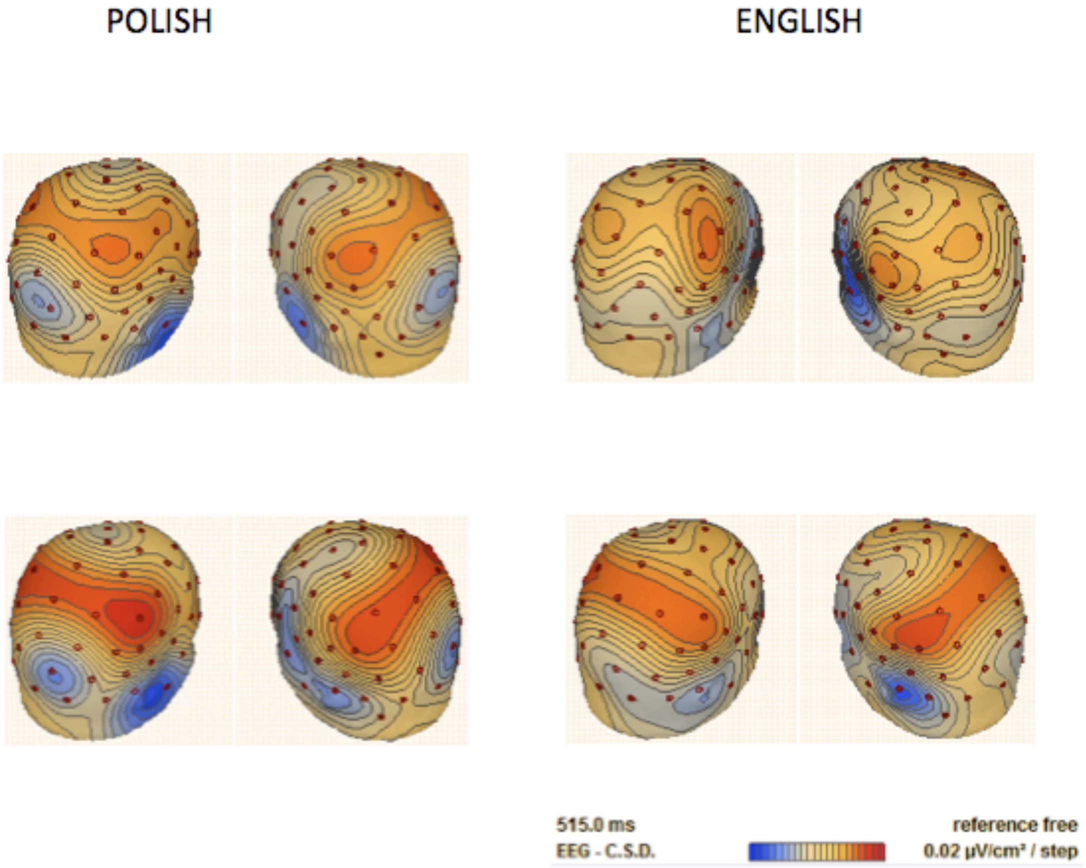


Figure 2.6 depicts the posterior parietal LPC in response to the *st* contrast (3-syllable) in the Polish (left) and English (right) listeners. A greater LPC for the different target (bottom) compared to the same target (top) occurred for both language groups

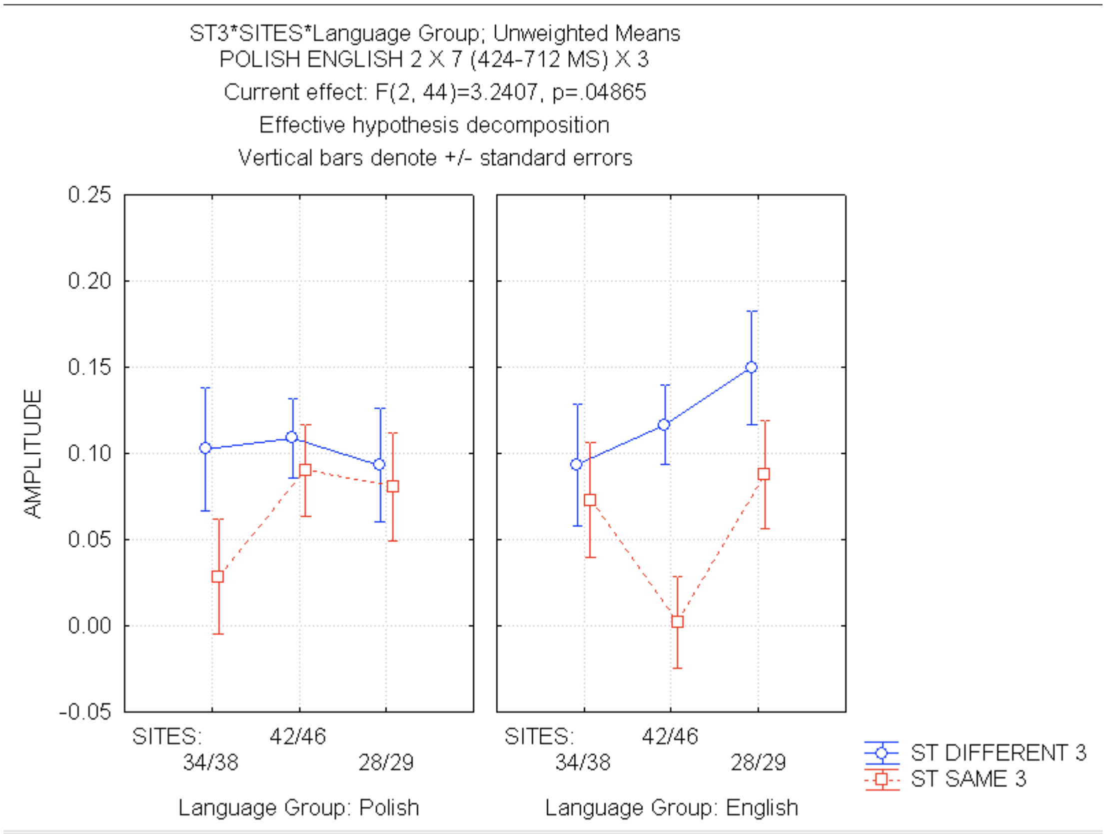


Figure 2.7 depicts topographical group differences with the English group having a larger response than the Polish group to the *st* contrast (3-syllable) from the averaged right-posterior parietal sites (42/46)

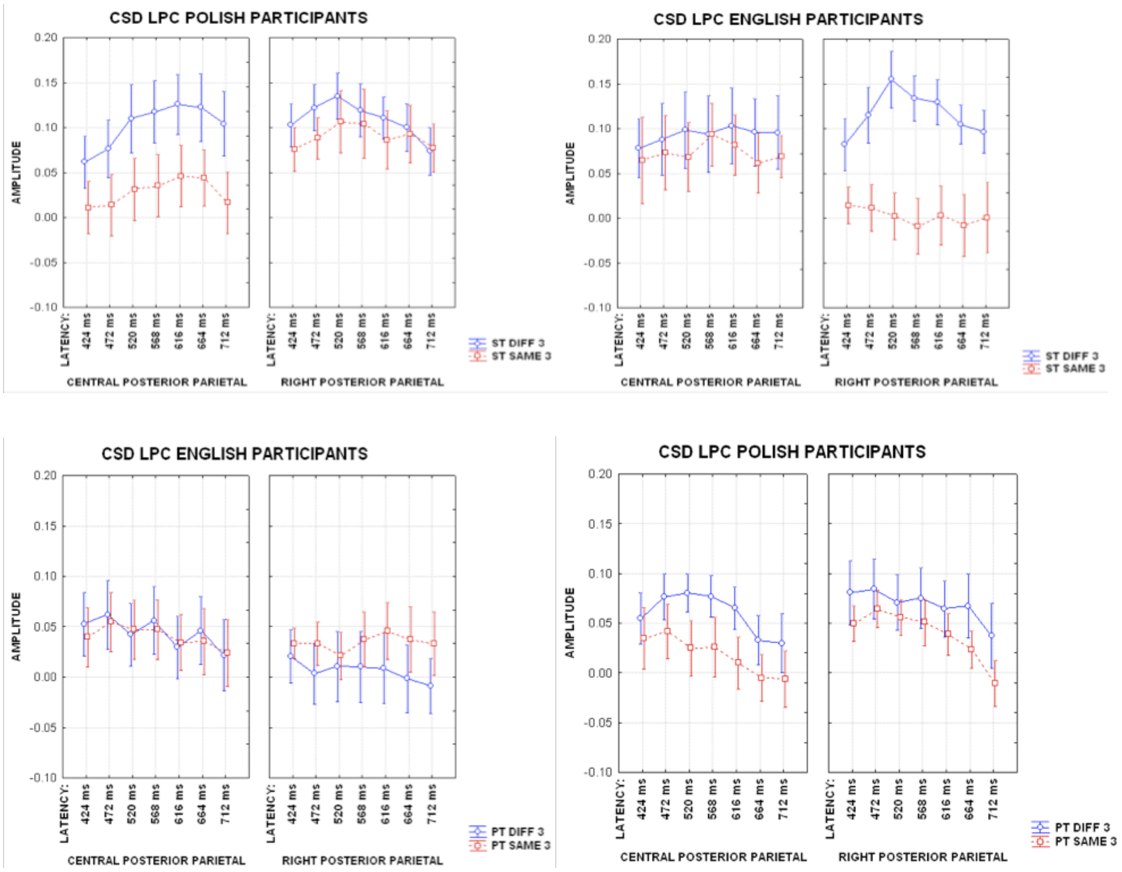
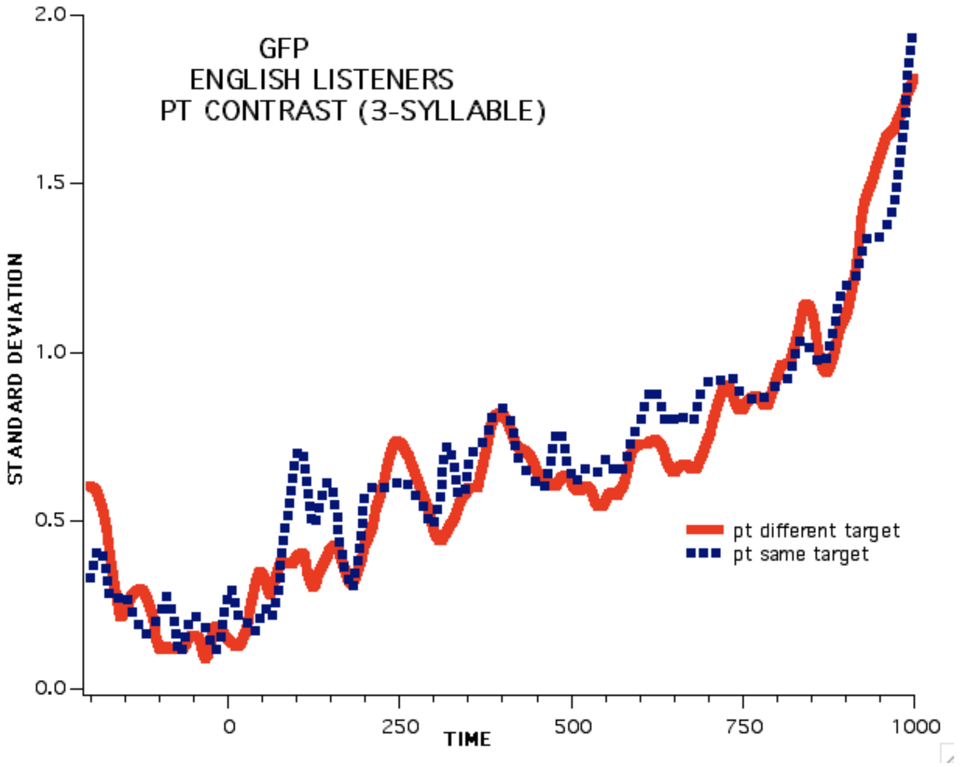


Figure 2.8 depicts the mean LPC response to *st* and *pt* contrasts (3-syllable) for the language groups. Notice the large response to the *st* different target from the right-posterior parietal sites for the English participants (top right) compared to the Polish participants (top left). Only in response to the pt contrast by the English participants is a reversed pattern evident (larger mean LPC to the same target) (bottom left)



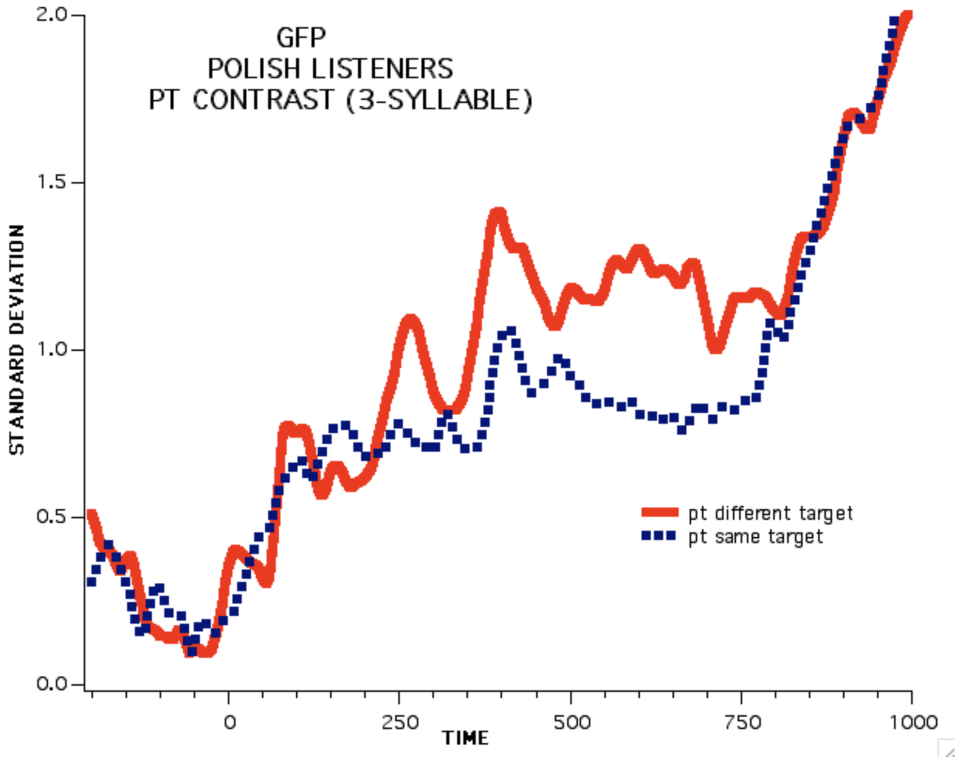


Figure 2.9 depicts the GFP analysis to the *pt* contrast (3-syllable) in English (top) and Polish (bottom) participants

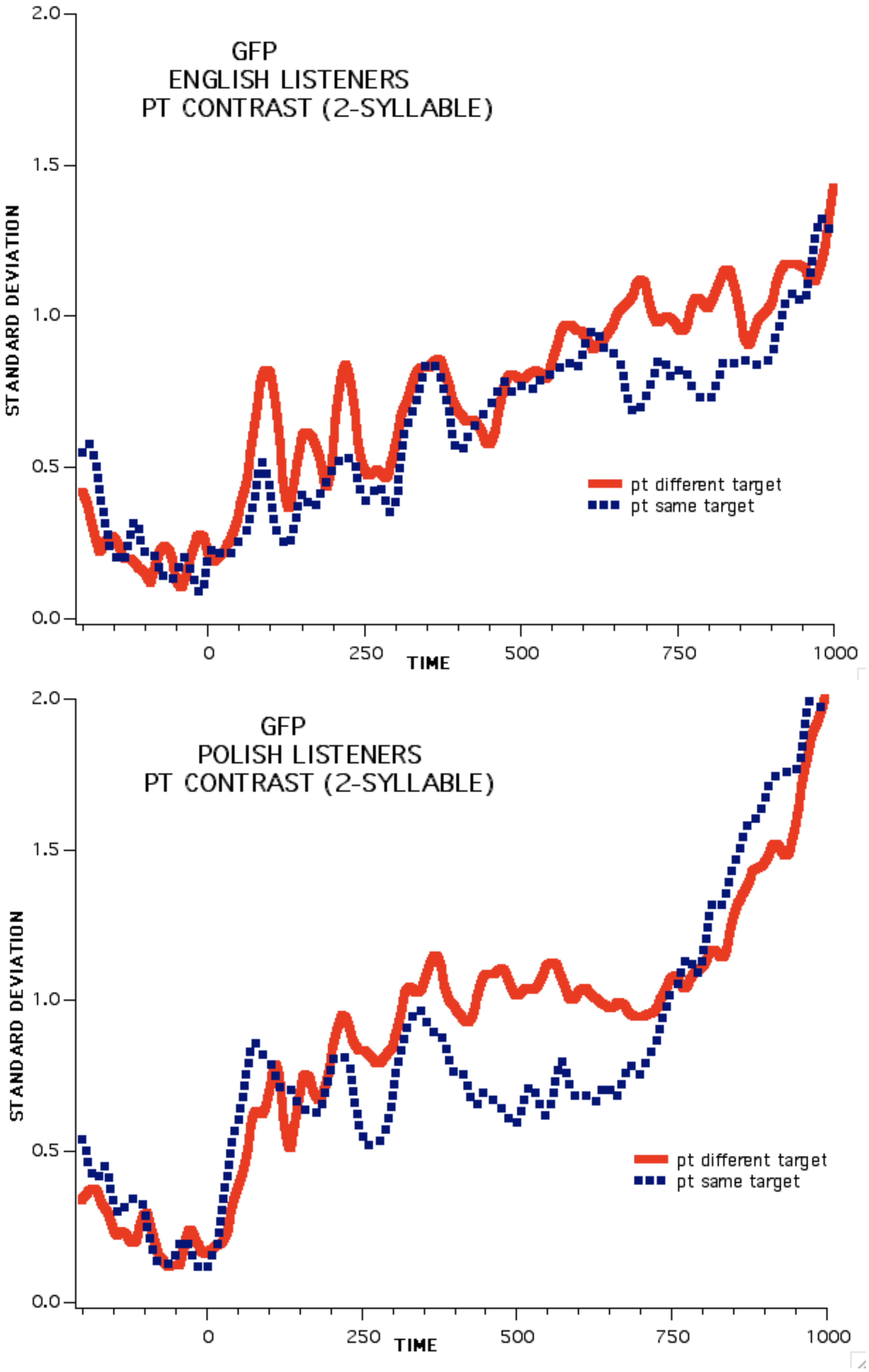


Figure 2.10 depicts the GFP analysis to the *pt* contrast (2-syllable) in English (top) and Polish (bottom) participants

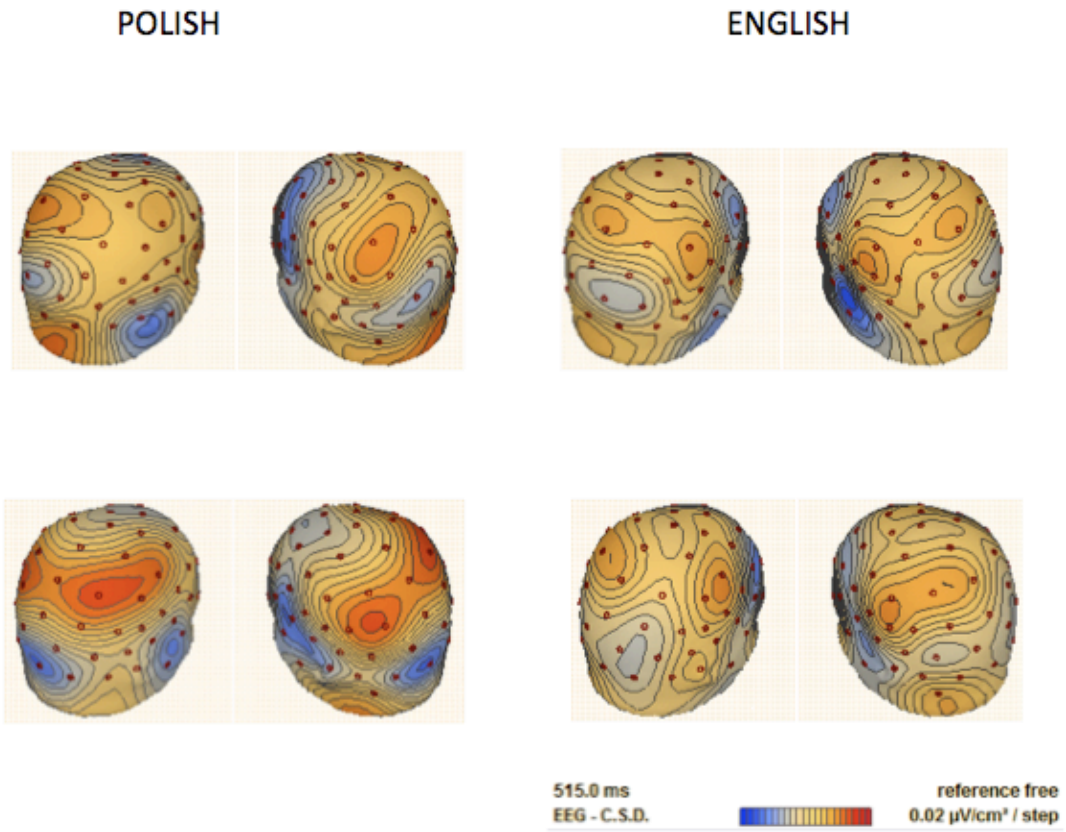


Figure 2.11 depicts the posterior parietal LPC in response to the *pt* contrast (3-syllable) in the Polish (left) and English (right) listeners. A greater LPC for the different target (bottom) compared to the same target (top) is evident for the Polish group but not the English group

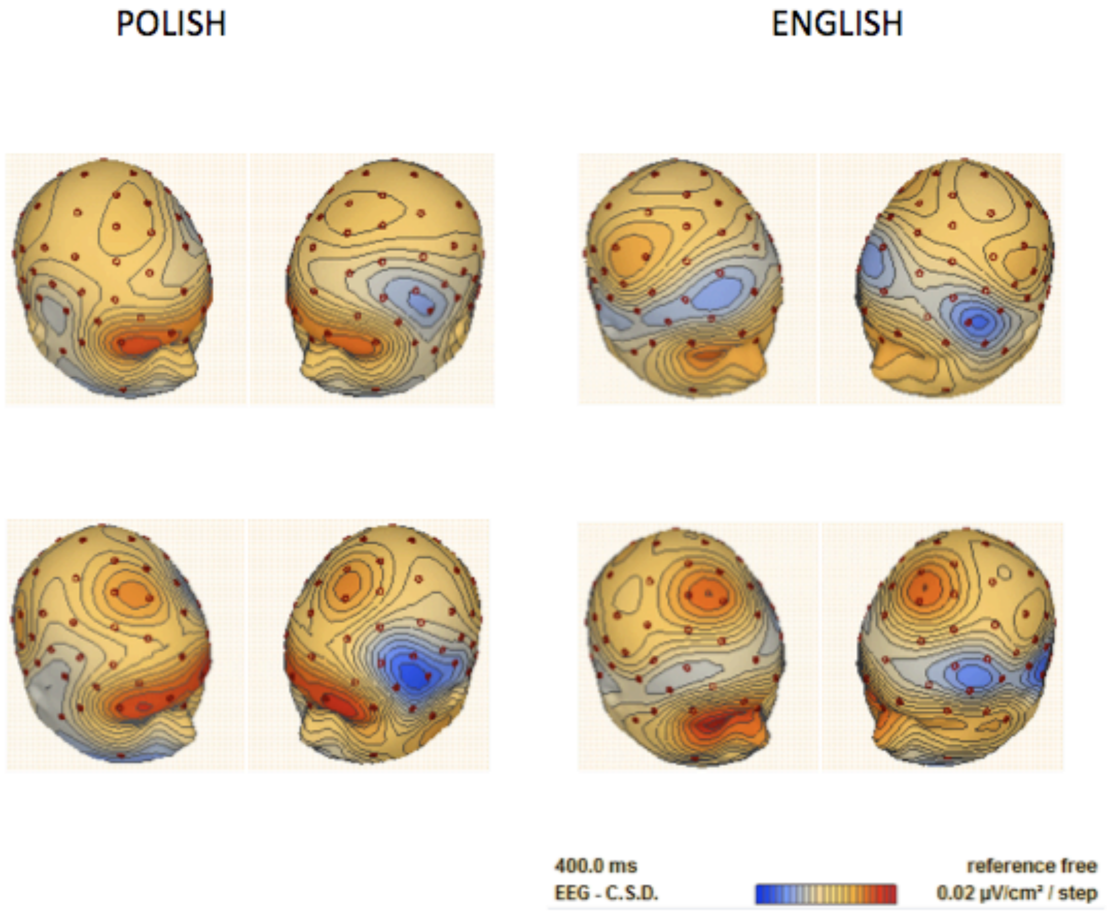


Figure 2.12 depicts the fronto-central P400 response to the *st* contrast (3-syllable) in the Polish (left) and English (right) language groups. The response to the different target (bottom) is greater than the response to the same target (top) in both groups of participants

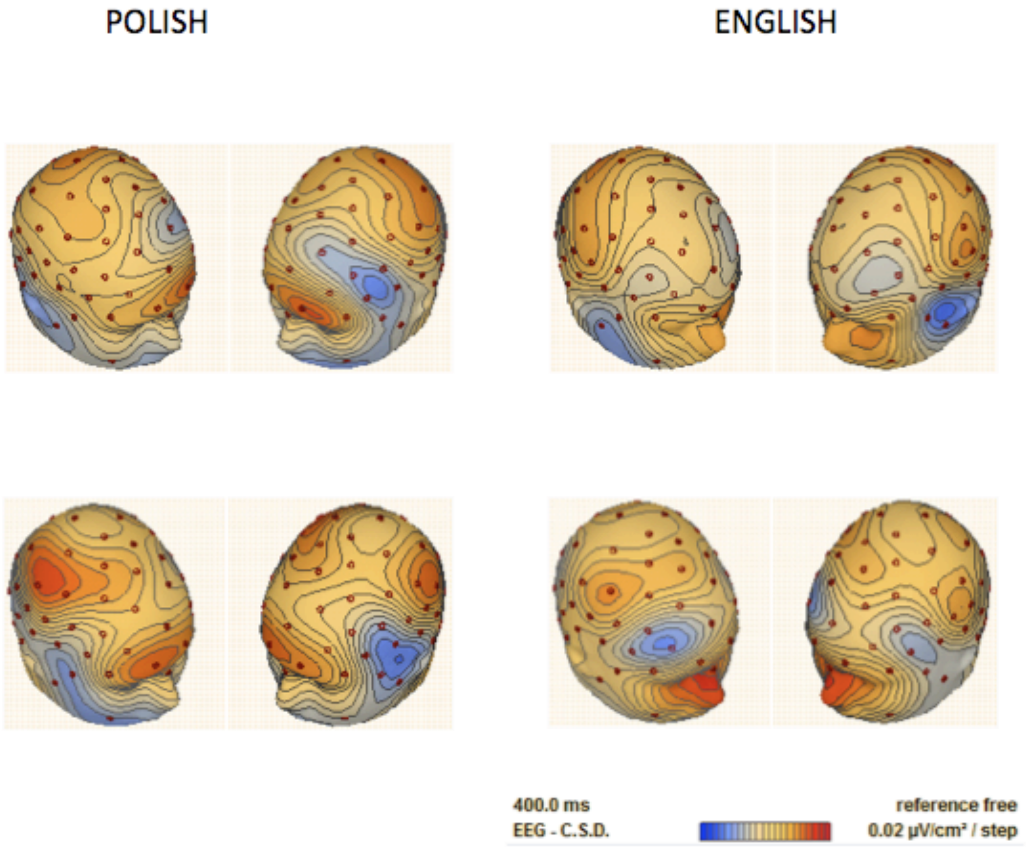


Figure 2.13 depicts the fronto-central P400 response to the *pt* contrast (3-syllable) in the Polish (left) and English (right) language groups. The response to the different target (bottom) is greater than the response to the same target (top) in both groups of participants

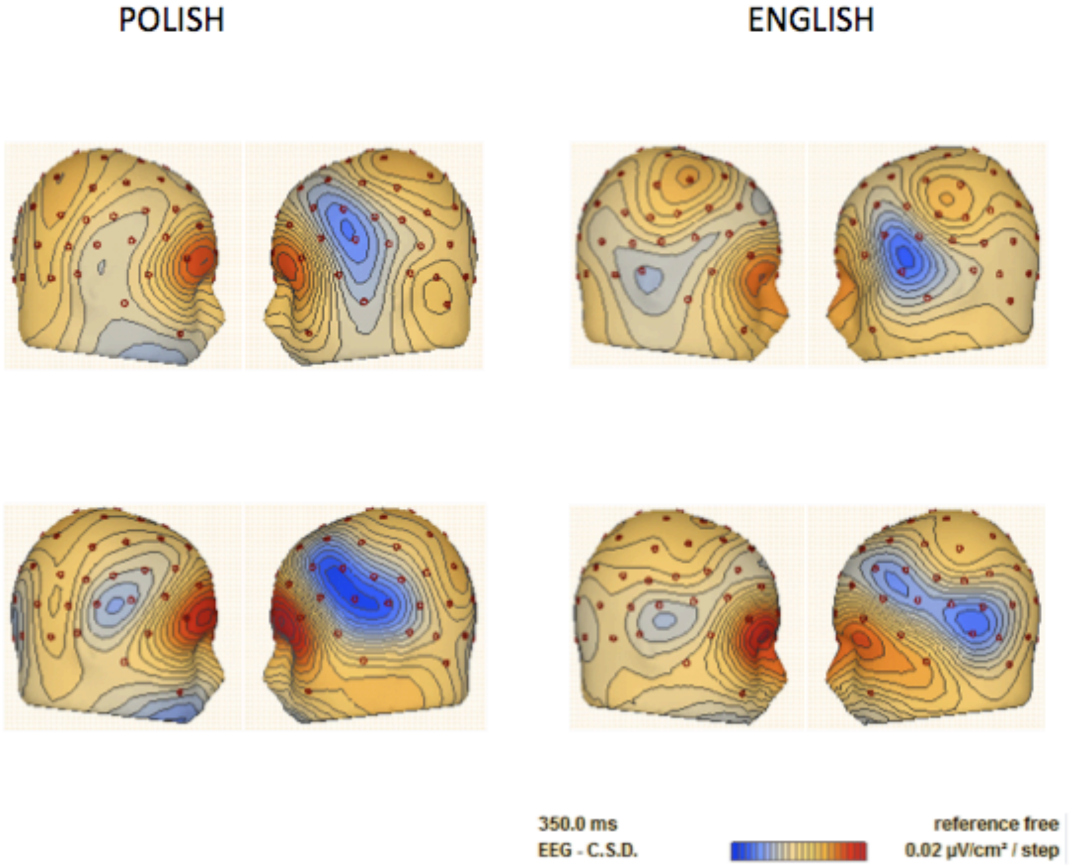


Figure 2.14 depicts the lateral temporal negativity in response to the *st* contrast (3-syllable) in the Polish (left) and English (right) language groups. The response to the different target (bottom) is greater than the response to the same target (top) in both groups of listeners

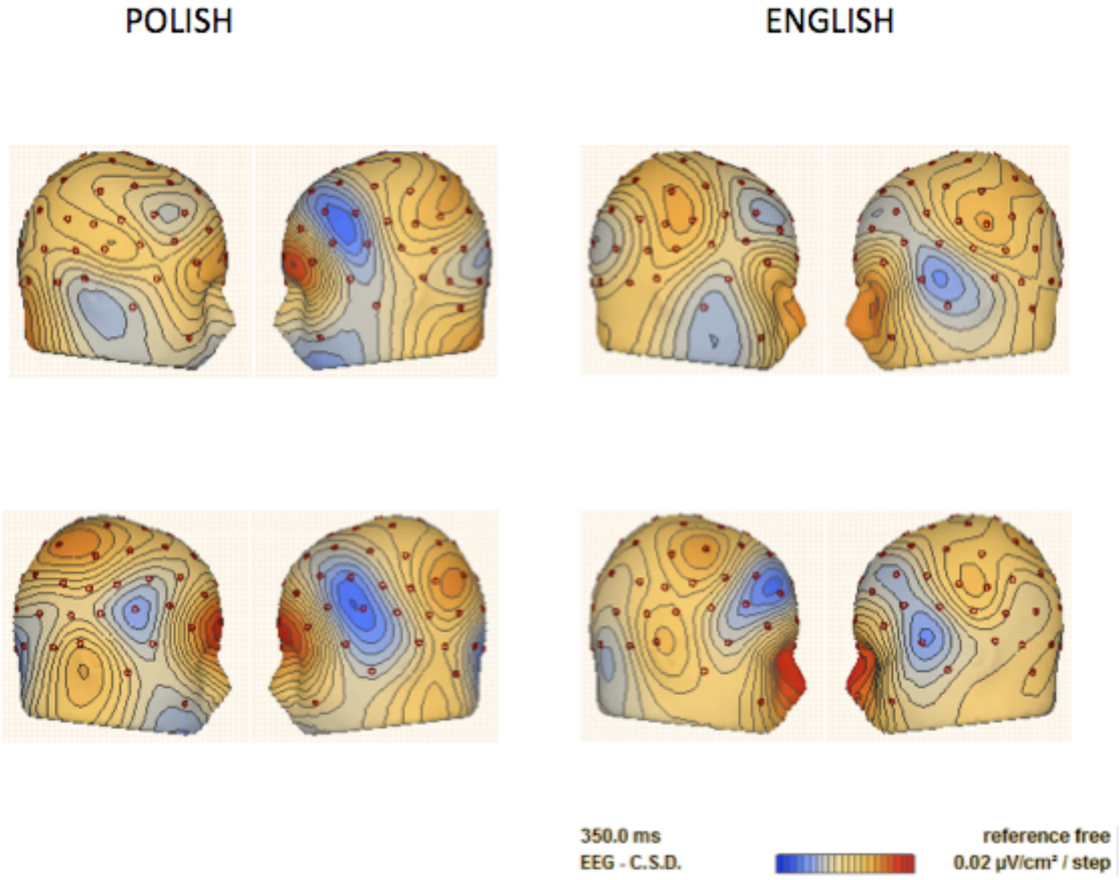


Figure 2.15 depicts the lateral temporal negativity in response to the *pt* contrast (3-syllable) in the Polish (left) and English (right) language groups. The response to the different target (bottom) is greater than the response to the same target (top) from left-posterior and right-anterior sites in both groups of participants

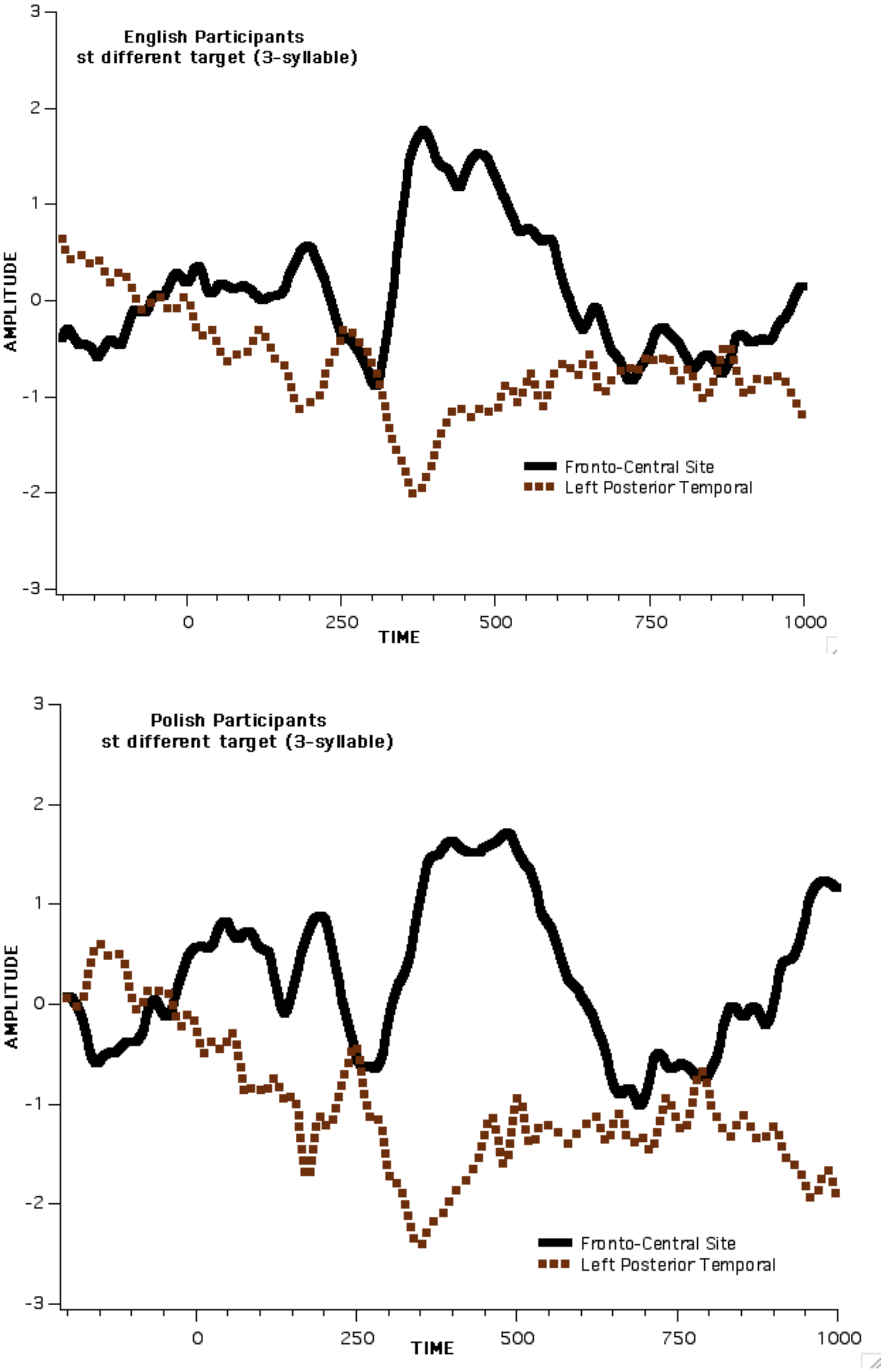


Figure 2.16 depicts the P400 waveform in response to the *st* different target and the inverse TN waveform for English (top) and Polish participants (bottom). Grand mean raw voltage data is shown

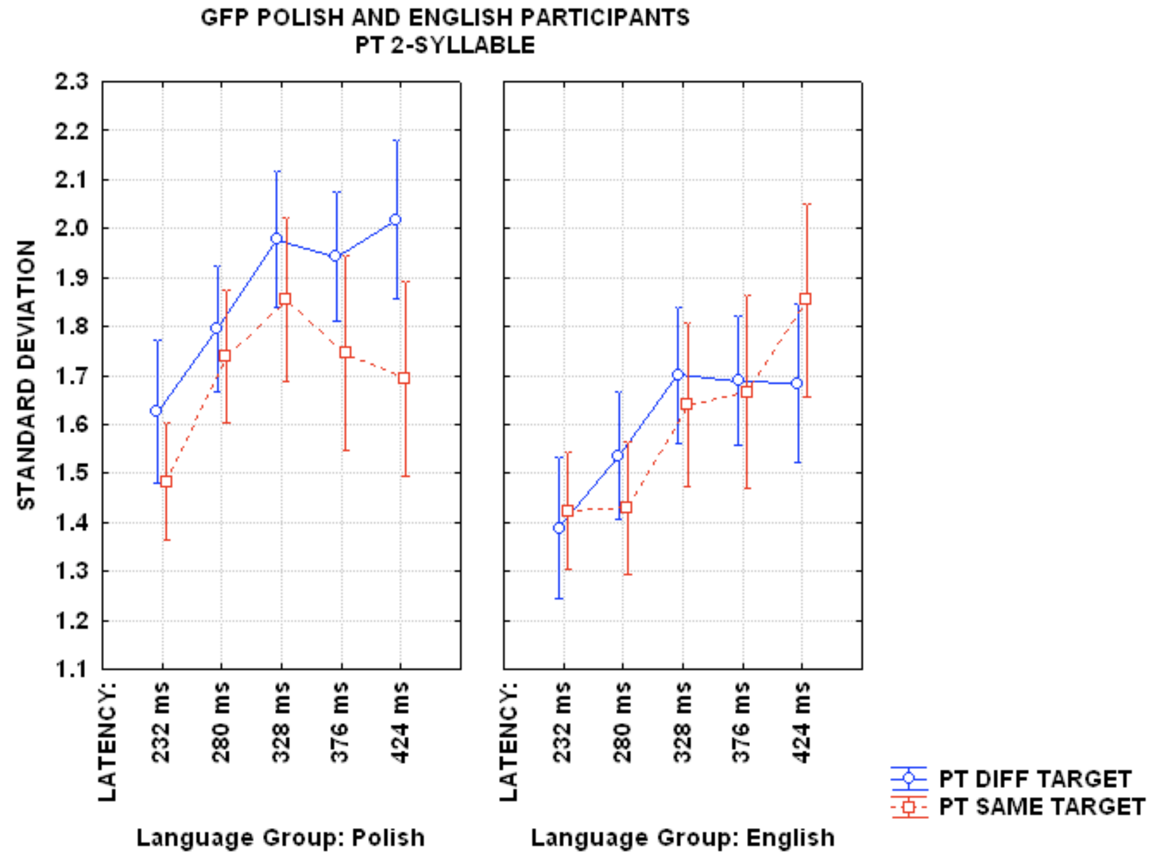


Figure 2.17 depicts an interaction of condition, time, and group showing a greater GFP for the different target (*pt* 2-syllable) relative to the same for the Polish participant data not seen in the English participant data

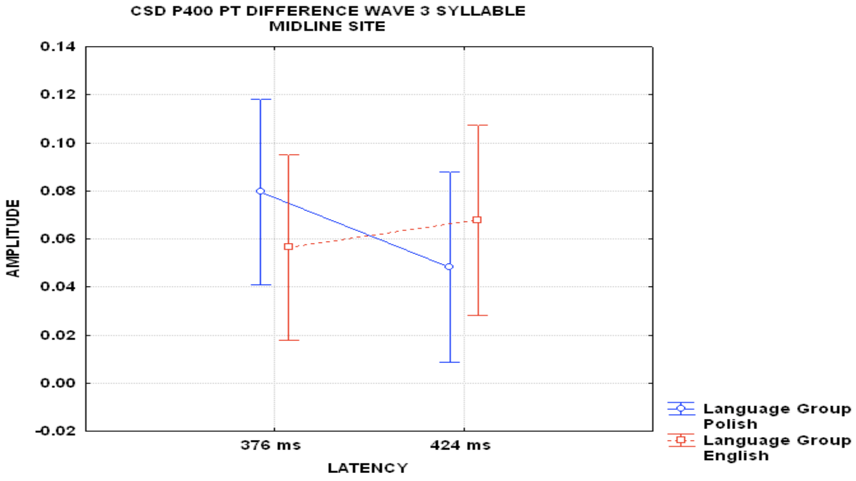


Figure 2.18 depicts the significant time x group interaction for the *pt* difference wave (3-syllable) at the midline fronto-central site revealing an earlier positive response for the Polish group

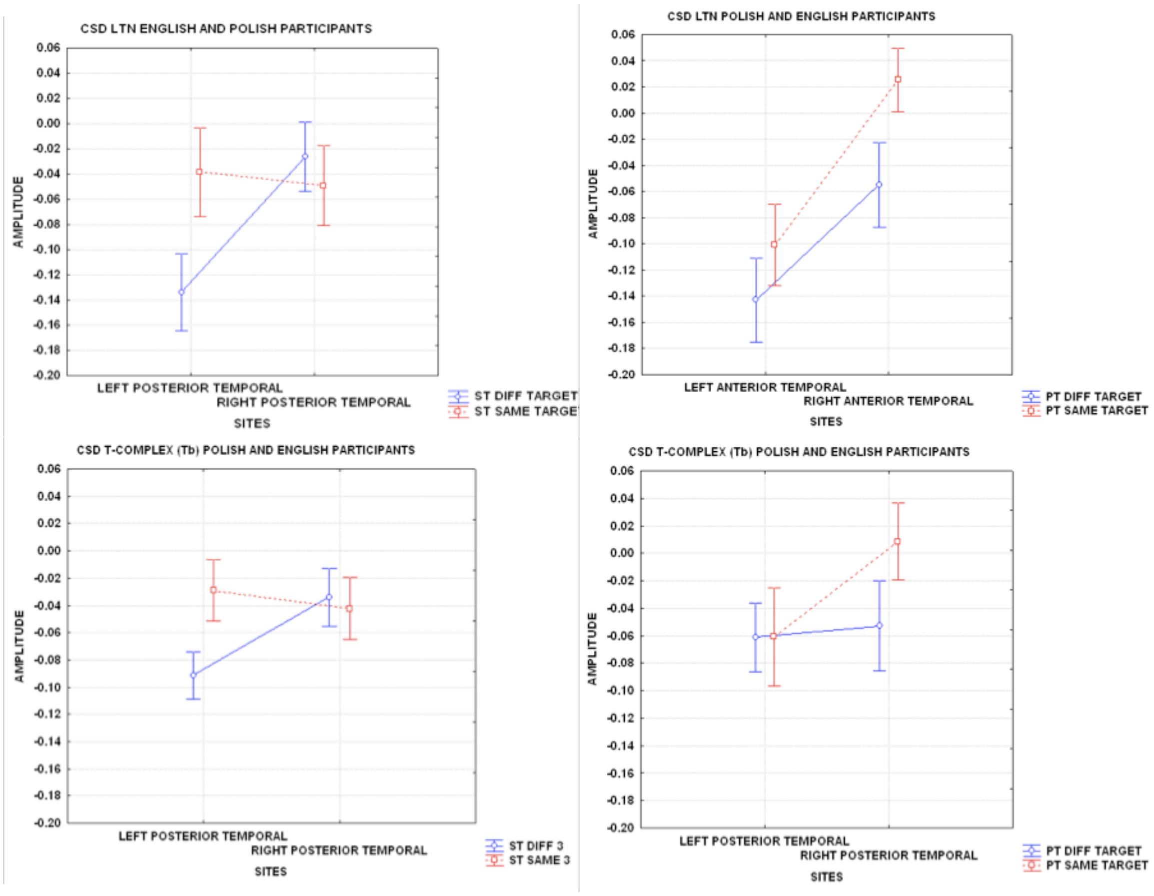


Figure 2.19 depicts the left-temporal hemisphere effect for the *st* contrast (328, 376 ms) (top left) not evident for the *pt* contrast (280, 328 ms) (top right). A left-temporal hemisphere effect at 184 ms (Tb component of the T-Complex) in response to the *st* contrast (bottom left) was not evident in response to the *pt* contrast (bottom right)

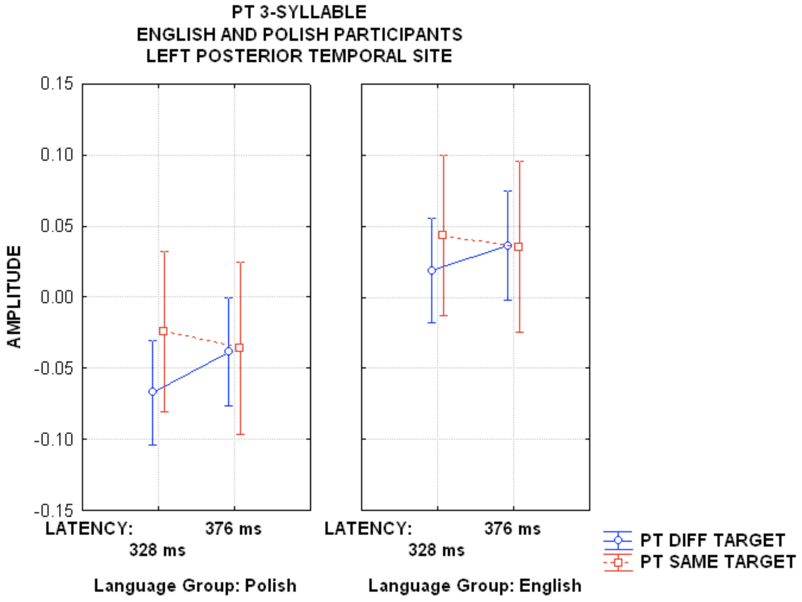


Figure 2.20 illustrates a greater negativity for the different target compared to the same at the 328 ms time interval from the left-posterior temporal site. The effect does not differ for the language groups

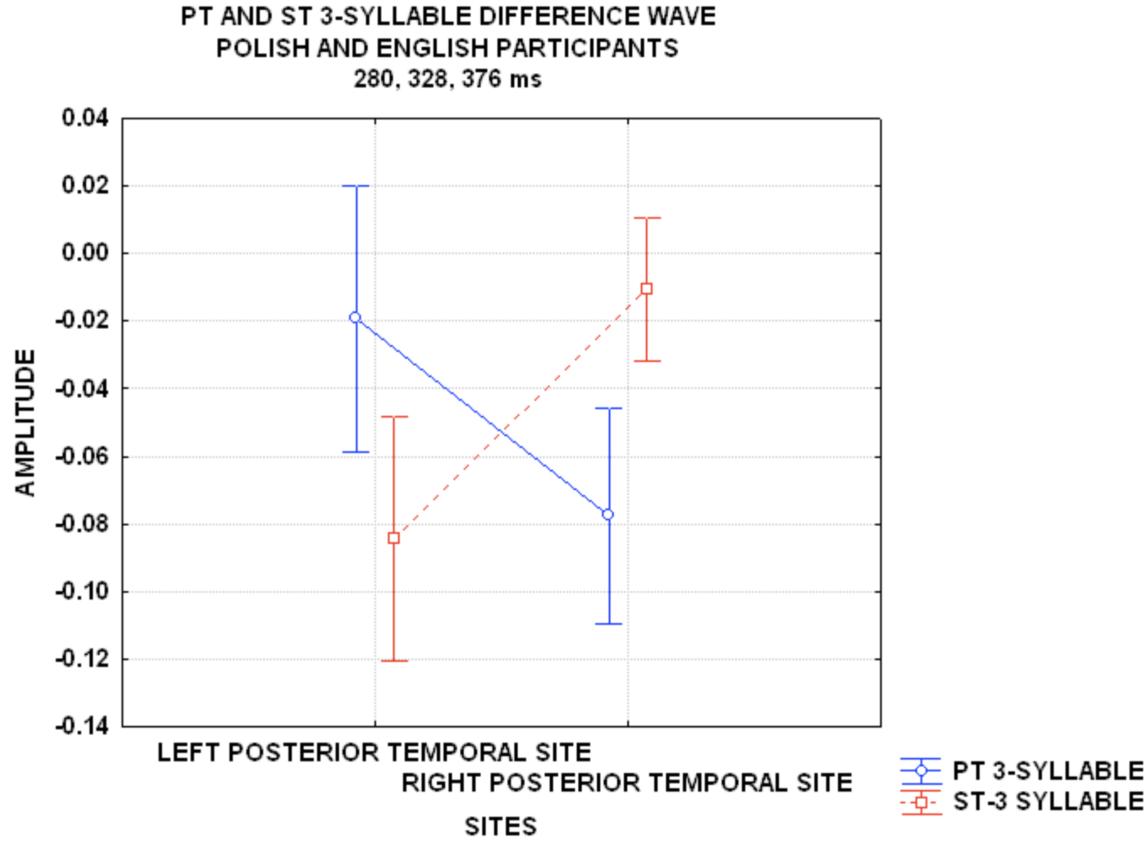


Figure 2.21 shows a relative left hemisphere effect for the *st* difference wave (3-syllable) and a relative right hemisphere effect for the *pt* difference wave (3-syllable) at posterior temporal sites

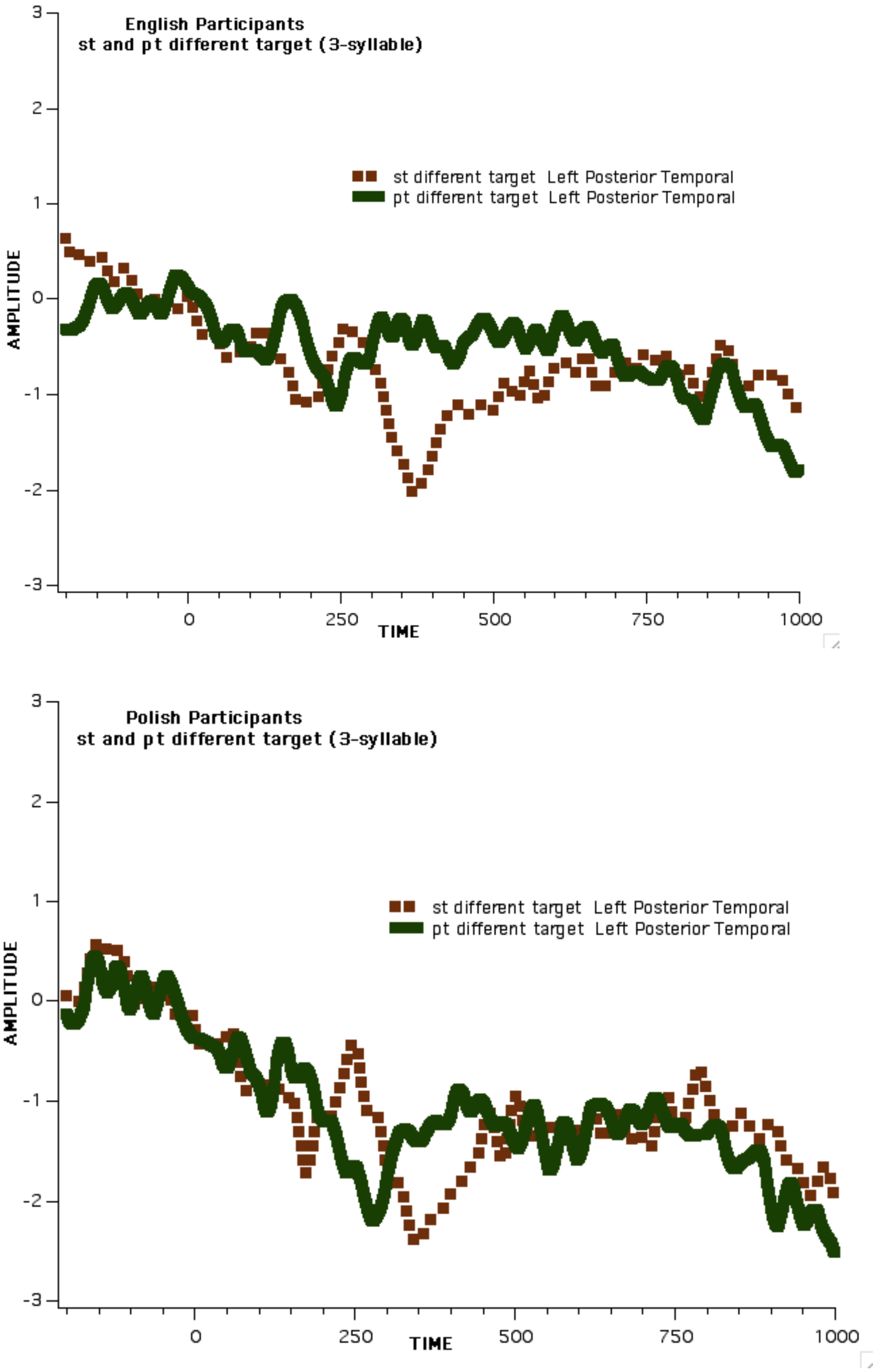
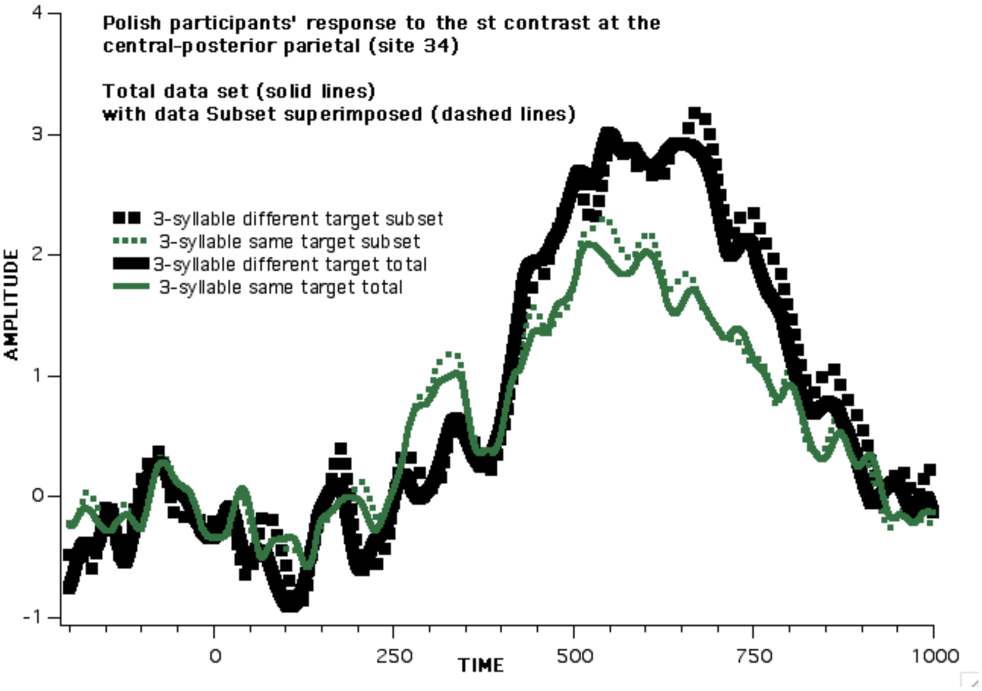


Figure 2.22 depicts two clear left temporal negative peaks (180 ms and 350 ms) in response to the *st* different targets that are less pronounced to the *pt* different targets. A similar response is evident in the English (top) and Polish participant (bottom) data. Grand mean raw voltage data is shown

APPENDIX A

A subset of ERP data was created by removing experimental trials containing 20 incorrectly perceived token nonsense words (*pt 3-syllable*) along with the 2-syllable *pt* nonsense words and 2 and 3 syllable *st*-nonsense words matched for rhyme from the ERP data set. For example, all trials containing the nonsense words *pətɪba* were removed along with trials containing *ptɪba, stɪba* and *sətɪba*. Thirty trials were excluded from each condition resulting in a remaining 70 trials. The procedure was applied to each individual's data set. In addition two real Polish words beginning with /st/, erroneously included as nonsense word stimuli, were among the eliminated trials. The subset of data created was graphed and compared to the total data set as shown in Appendix A Figures A.2.1 through A.2.4.



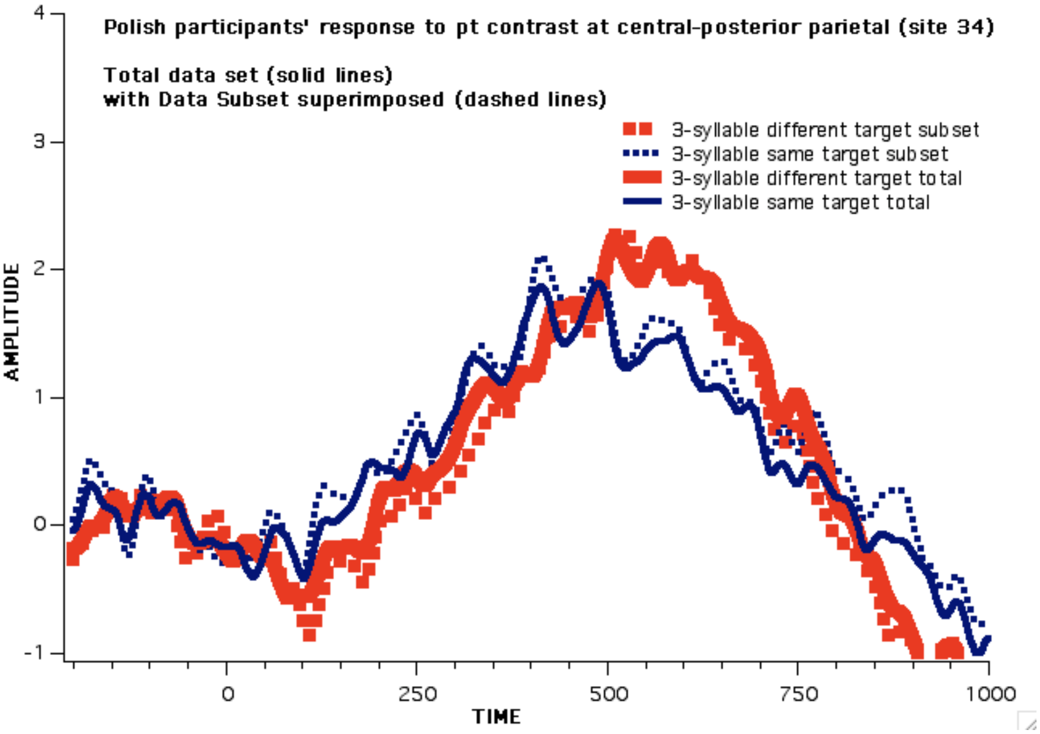


Figure A.2.1 depicts the Polish participants' LPC response to the *st* contrast (top) and the *pt* contrast (bottom) from the central-posterior parietal site (site 34). The data subset (dashed lines) graphed is superimposed on the total data set (solid lines)

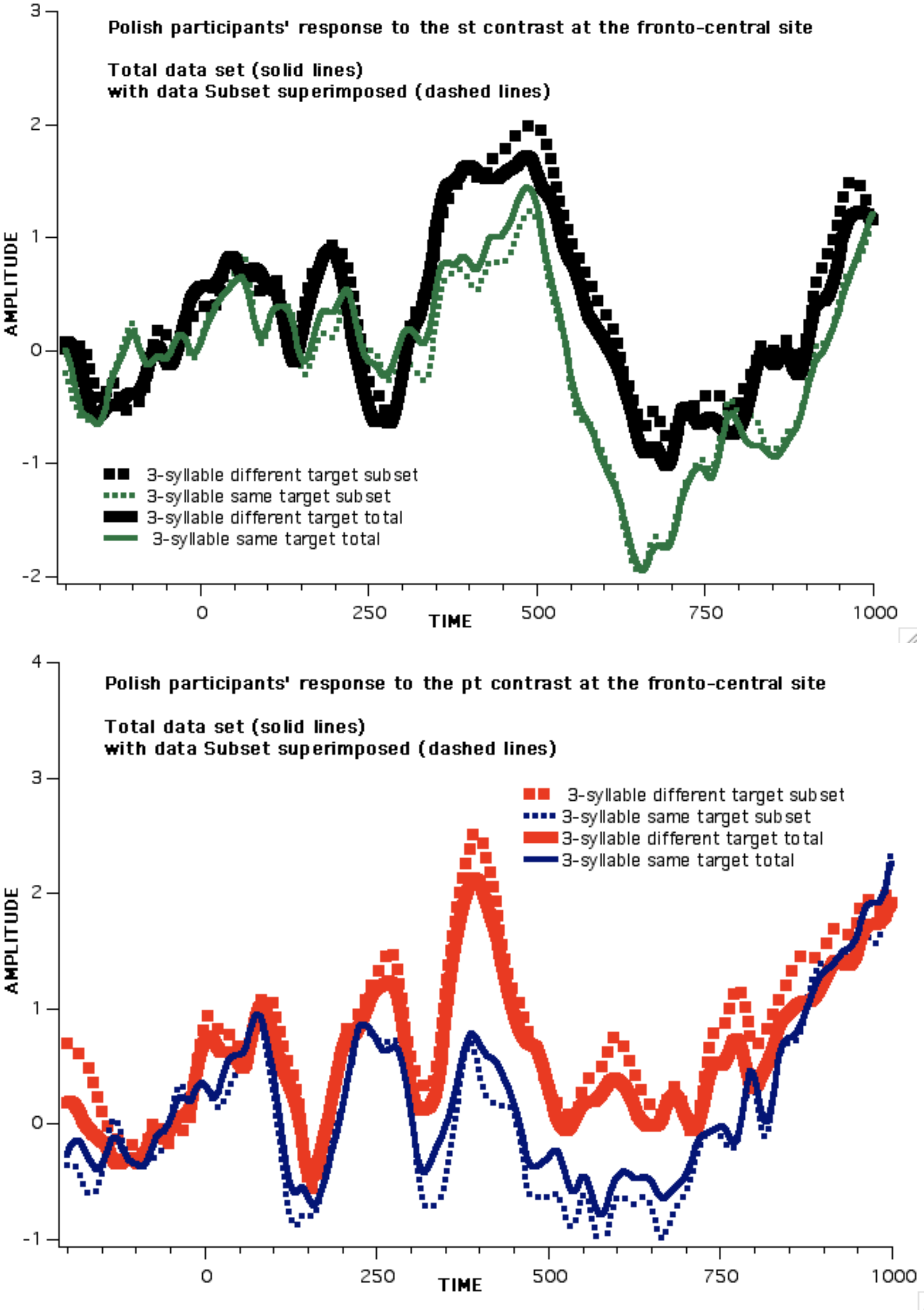
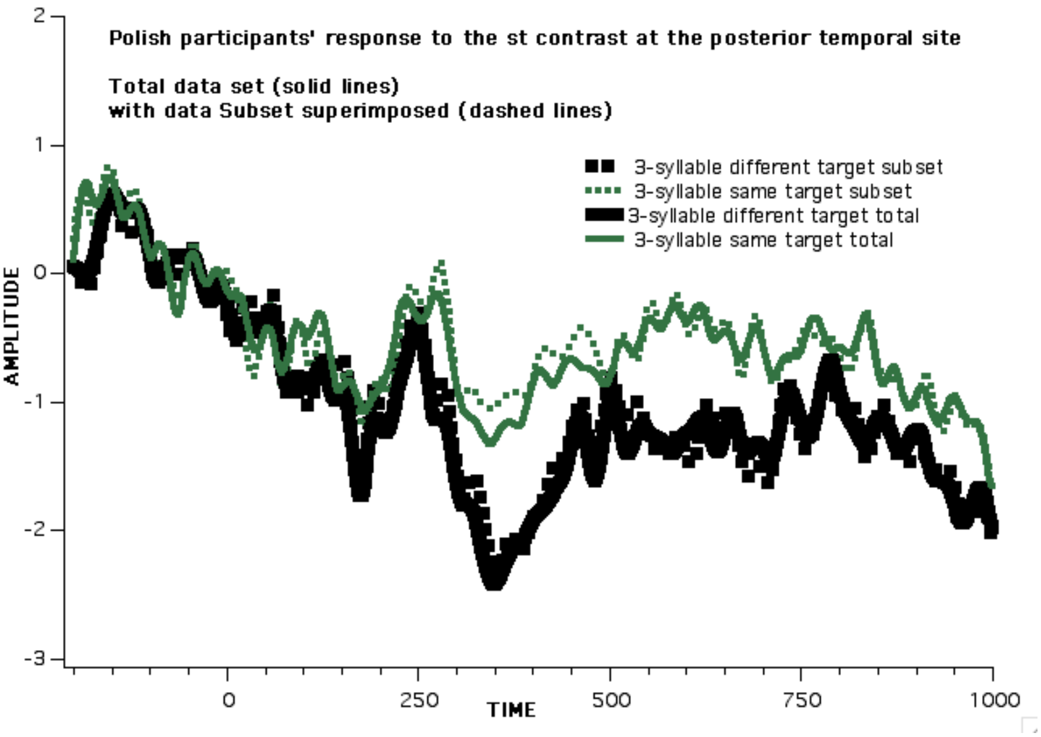


Figure A.2.2 depicts the Polish participants' P400 response to the *st* contrast (top) and the *pt* contrast (bottom) from the midline fronto-central site. The data subset (dashed lines) graphed is superimposed on the total data set (solid lines)



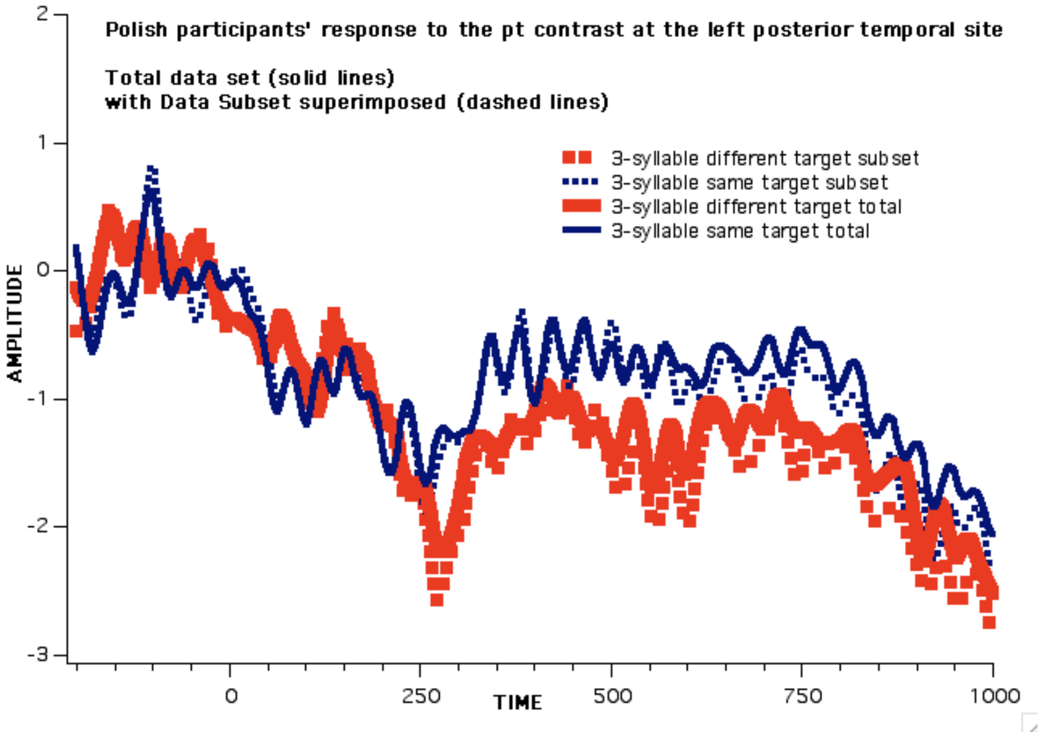
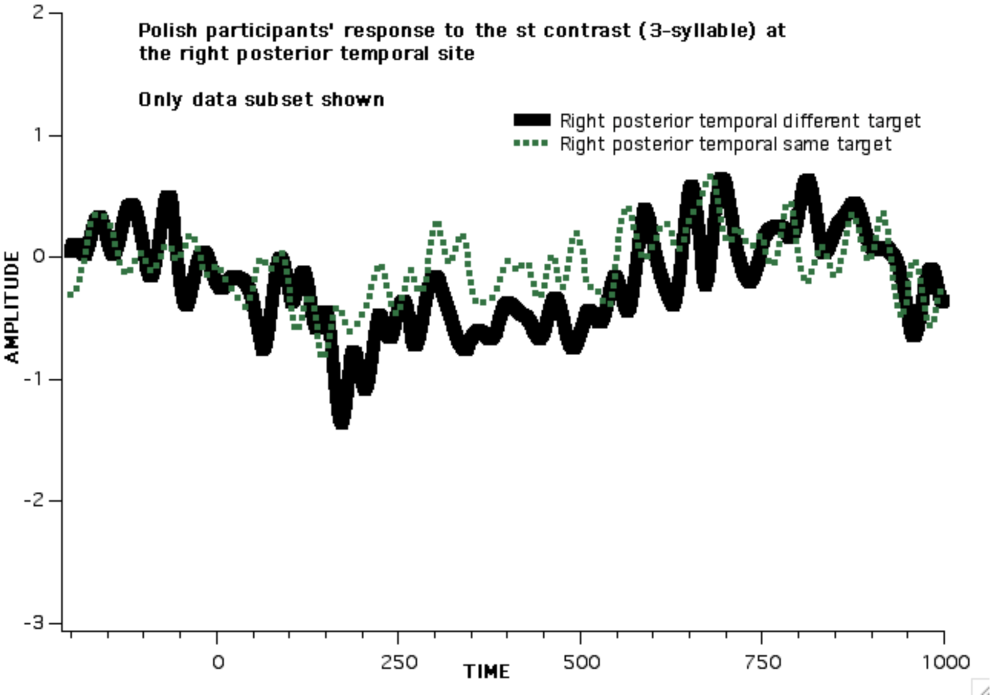
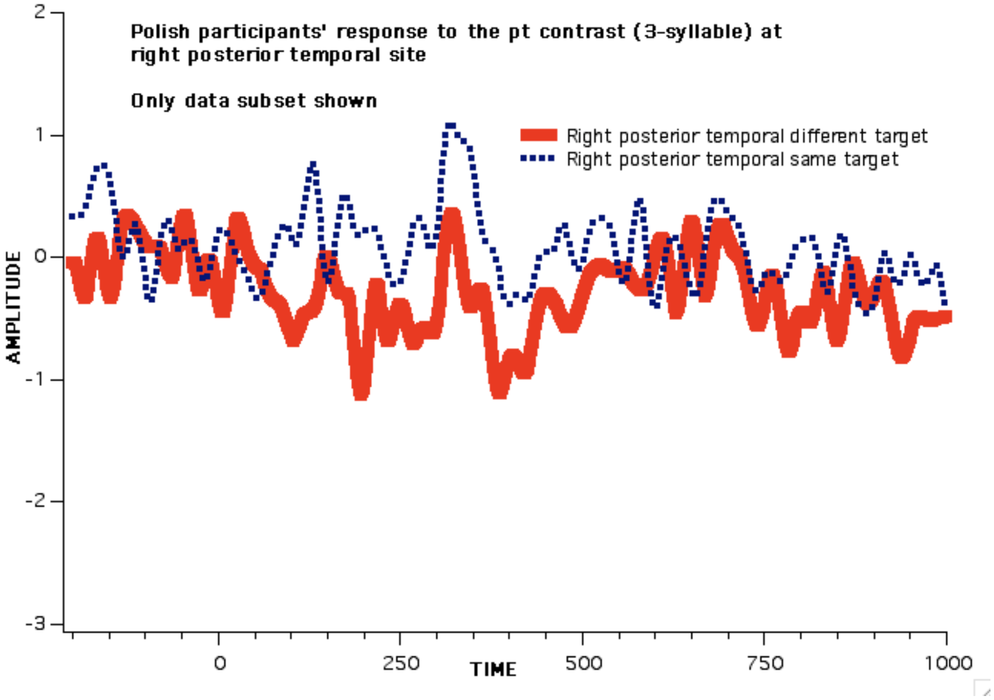


Figure A.2.3 depicts the Polish participants' ERP response at the left posterior temporal site to the *st* contrast (top) and the *pt* contrast (bottom). The data subset (dashed lines) graphed is superimposed on the total data set (solid lines)





A.2.4 depicts the Polish participants' ERP response at the right posterior temporal site to the *st* contrast (top) and *pt* contrast (bottom). Only the data subset is shown

CHAPTER 3

THE P1-N1-P2 AND THE T-COMPLEX IN NATIVE-ENGLISH AND NATIVE-POLISH LISTENERS: IS NATIVE-LANGUAGE SPEECH PERCEPTION REFLECTED IN THE SENSORY-OBLIGATORY ERP COMPONENTS?

INTRODUCTION

A major issue of our time, in the neuroscience of speech and language, concerns the manner in which the acoustic signals of speech are translated in cortical auditory pathways into integrated neural signals necessary for speech perception. Comparison of native-language speech perception and neurophysiological indices of speech processing permit an exploration of the effects of varying auditory-phonetic input from the natural environment on stages of neural speech processing. The current research explores the effects of native-language experience on early stages of cortical speech processing.

The investigation of the neural correlates of native-language perception of words has focused on relatively late stages of cortical speech processing, such as the N400 and the Late Positive Component (LPC). These studies have revealed that experience with the phonological patterns of one's native-language influences relatively late stages of neural processing (Dehaene-Lambertz, Dupoux & Gout, 2000; Friedrich, Schild, & Roder, 2009). One question that has not been addressed, however, is to what extent does experience with the phonological patterns of a listener's native-language influence earlier levels of processing.

The P1N1P2 complex demonstrates access to the acoustic cues within the speech signal necessary for speech perception and production (Martin, Tremblay & Korczak, 2008) and indexes earlier levels of processing than the N400 and the LPC.

The P1N1P2 complex is a sensory-obligatory detection response to the onset and offset of auditory stimuli at midline central electrode sites in adult subjects. The acoustic change complex is the P1N1P2 response to the changing acoustic features or sounds within a word, which overlap resulting in a signature waveform for a word (Martin, Tremblay & Stapellis, 2007; Kaukoranta, Hari & Lounasmaa, 1987). The signature waveform in response to the word "say" /sei/ has been shown to be a combination of the P1N1P2 complex in response to the sound /s/ alone and the vowel /ei/ alone (Ostroff, Martin, and Boothroyd, 1998).

Auditory cortex on the superior temporal plane is the primary contributor to the P1N1P2 complex recorded from fronto-central electrode sites. N1 generators, also, include lateral surface auditory cortex and, possibly, frontal motor or premotor cortex. The P2 waveform has multiple generators including primary, secondary auditory cortex and the reticular activating system. While the P1, N1 and P2 waveforms reflect onset detection of auditory stimuli they each have differing underlying neural sources and independent response patterns (Steinschneider & Dunn, 2002; Martin, Tremblay, & Korczak, 2008; Naatanen & Picton, 1987; Ceponiene, Alku, Westerfield, Torki & Townsend, 2005). For example, research has demonstrated that N1 and P2 amplitude and latency decrease with increases in frequency of the acoustic signal (Crowley & Colrain, 2004), however, Wunderlich and Cone-Wesson (2001) found that latency of the N1 component decreases with increased frequency but the latency of the P2 component did not. This suggests that the P1N1P2 complex to speech stimuli will reflect varied functions.

The T-complex is a sequence of auditory evoked potentials recorded from temporal sites and is sensitive to acoustic properties of stimuli. The peaks consist of a negative wave at approximately 70 to 80 ms (i.e., Na), a positive wave at approximately 100 ms (Ta) and a large negativity around 140-160 ms (Tb). These components are best obtained from lateral temporal electrode sites because they show a radial orientation from their source (Tonnquist-Uhlen, Ponton, Eggermont, Kwong, & Don, 2003; Wolpaw & Penry, 1975; Friedrich, Schild, & Roder, 2009; Naatanen & Picton, 1987). Dipole source modeling suggests that the T-complex originates from secondary auditory cortex on the lateral surface of the superior temporal gyrus and is at least partially independent from P1N1P2 complex (Ponton, Eggermont, Khosla, Kwong, & Don, 2002).

Differing N1, P2 and T-Complex response patterns have been found in special populations leading researchers to investigate the effect of auditory perception training on these components. For instance, trained musicians were found to have a larger P2 (fronto-central) and a larger left temporal negativity at approximately 150 ms than non-musicians in response to musical tones. This finding suggests an influence of auditory perceptual skill on the sensory-obligatory components (Shahin, Bosnyak, Trainor & Roberts, 2003). The results of auditory discrimination training studies employed to determine the relationship between speech perception and the P2 waveform have shown contradictory results (Tremblay & Kraus, 2002; Sheenan, McArthur & Bishop, 2005, Reinke, He, Wang & Alain, 2003). In one study, pre and post training behavior and ERP measures were compared in American participants taught to identify syllables having a non-native prevoicing/voicing distinction of 10 ms (i.e., -20ms VOT labeled “mba” or -10ms VOT labeled “ba”). Identification of the syllables containing the VOT distinction improved in 5 out of 7 participants. The ERP responses to the “mba” or “ba” syllables did not differ (pre or post-training) and were, therefore, combined. The P1N1P2 post-training response to the combined syllables revealed amplitude changes with the greatest change occurring for the P2 component. Examination of individual response patterns, however, revealed that post-training modifications in ERP patterns did not correlate with modifications in individuals' behavioral response. The researchers argued that the training process activated an alerting mechanism that may support perceptual learning resulting in the P2 response. In another behavioral and ERP study, participants were trained to identify "double vowels" within vowel pairs. The vowels were created by combining the digital waveforms of two different vowels. A follow-up behavioral vowel identification task found the trained group to have improved significantly and the untrained-control group to have improved marginally. Post-training ERP results revealed a significant decrease in P2 latency at right-temporal sites and a significant increase of P2 amplitude at left-temporal sites for the trained group only. However, both the trained and untrained groups showed modifications in N1 latency at central sites and P2 amplitude at central sites and right-temporal sites. Modifications in N1 and P2 components at central sites by the untrained group occurred after only brief exposure drawn to the stimuli during the pre-testing session (one week prior) suggesting that attention to stimuli affects the components. Also, no attempt was made to correlate individuals' behavioral performance with ERP waveform changes and, therefore, ERP modifications might have reflected greater attention directed to experimental stimuli by the trained group as a result of four days of experience with the experimental stimuli (Reinke, He, Wang & Alain, 2003). Attention to experimental stimuli clearly influences ERP components after ~50 ms and has been shown to affect N1, P2 components and P3 components (Eggermont, 2007; Picton & Hillyard, 1974; Steinschneider & Dunn, 2002). The aforementioned studies suggest that experience with acoustic signals for long durations (i.e., musicians versus non-musicians) as well as brief exposure to acoustic features of sounds (including a VOT non-native sound contrast) affects the sensory-obligatory components. Presently, we compare the effect of exposure or lack of exposure to acoustic features of sounds during native-language development on the P1N1P2 complex and the T-Complex complex.

Naatanen & Winkler (1999) addressed the issue of speech perception and its relationship to the stage of processing reflected in the N1 component stating that the N1 component cannot reflect integrated sound because the response does not meet the requirement "completeness" (p 833). They write that the supratemporal N1 response mainly reflects transient onset and offset detection and does not code temporal information of sound unless a change in the auditory signal occurs. They propose that the N1 component may reflect either the final stage of auditory processing before sound representation (reflected within the mismatch negativity, MMN) or N1 generators (from fronto-central sites) may not lie within a serial path with generators involved in sound representation as suggested by Davis & Zerlin (1966). The latency of the MMN, a component that can index speech discrimination, overlaps with the latency of the P2 component. For this reason, it has been suggested that the P2 component may reflect sensory encoding leading to perception or, perhaps, mechanisms to facilitate or inhibit conscious perception (Naatanen & Winkler, 1999; Ceponiene et al., 2005; Tremblay & Kraus, 2002). Heightened awareness of auditory stimuli following exposure, evident through P2 modifications in the aforementioned studies, suggests that P2, in part, reflects neural activity facilitating or inhibiting conscious perception (Ceponiene et al., 2005).

The current research was part of a larger project, which examined contextual effects on perception of /pt/, a consonant cluster that occurs in English, but, not in the context examined, word onset. The acoustic features of a phoneme (e.g., /p/) vary depending on the context of that phoneme, that is, the specific sound, which precedes or follows, or the presence or absence of a period of silence preceding or following. As an example, if /p/ is heard in syllable onset (e.g., “pat”), the phone consists of a rapid *rise* in the first formant frequency during opening of the lips. If /p/, on the other hand, is heard in syllable coda (e.g., “tap”), the phone consists of a rapid *fall* in the first formant frequency, related to lip closure for the /p/ (Bordon, Harris, & Raphael, 2003 p 113). The phonotactic rules of a language (i.e., the permissible sequences of phonemes, including consonant clusters in onset and coda positions) determine all the allowable contexts of a phoneme. Both native-Polish and native-English listeners experience the /pt/ cluster in their language (e.g., Polish: “ptak”- bird, “Petronella”- a fairy tale character; English; “kept”, “except”) but only the Polish group experiences the cluster in word onset. Hence, the Polish group, alone, experiences the acoustic features of the /pt/ consonant cluster in the context, word onset (see Chapter 2 for further detail).

Perception of the /pt/ cluster was explored within a sound contrast, /pt/ versus /pət/, within nonsense word stimuli in native-English and native-Polish listeners because the sound contrast is a linguistic contrast in word onset in Polish (e.g., Polish: “ptak”- bird, “Petronella”- a fairy tale character) but is not a linguistic contrast in word onset in English (e.g., English: “petunia”). The contrast /st/ versus /sət/ within nonsense words was presented as a control because both languages have the contrast in word onset (e.g., Polish: “stal”- steel, “suterena”- basement; English: “step”, “sateen”). The participants were presented with same and different nonsense word pairs (e.g., /ptimɑ-ptimɑ/, /pətimɑ-pətimɑ/, /ptimɑ-pətimɑ/, /pətimɑ-ptimɑ/, /stimɑ-stimɑ/, /sətimɑ-sətimɑ/, /stimɑ-sətimɑ/, /sətimɑ-stimɑ/) and were required to determine if the second word in the pair had two or three syllables. ERPs were recorded and epochs were time-locked to both the first and second words in the pair. As reported in Chapter 2, the behavioral task revealed that the Polish listeners were able to distinguish the two and three-syllable nonsense words that began with /pt/ or /pət/ and English listeners were not. Also, the language groups' behavioral pattern of response was reflected in a late positive component of the ERP waveform having a posterior parietal topography peaking at approximately 500 ms and, to some extent, in a positive focal fronto-central component peaking at 400 ms (see Chapter 2 for details). In this chapter, we report the P1N1P2 and T-Complex responses to the first word in the word pairs to determine whether the native-language perceptual behavior is reflected in early event-related potentials, indexing acoustic processing.

We hypothesized that native-language speech perception patterns would not be reflected at the early stage of cortical processing indexed by the P1N1P2 complex and the T-Complex, rather, universal acoustic properties would be indexed at this stage. Specifically, we predicted different auditory signature waveforms at fronto-central electrode sites for the four word forms presented, 2-syllable *pt* onset words, 3-syllable *pt* onset words, 2-syllable *st* onset words and 3-syllable *st* onset words and we predicted that these waveforms would be the same for both native-language groups. Furthermore, we predicted the T-Complex response to the four word types would be the same for both groups.

METHOD

Twenty-four subjects participated in the study (English: 8 female and 4 male; mean = 29 years; Polish: 8 female and 4 male; mean = 30 years) (See Chapter 2, method section). The native-Polish listeners were bilingual Polish-English speakers who had emigrated from Poland to the United States as young adults, after 15 years of age. The English listeners reported having no exposure to Polish or other Slavic languages. Participants had no history of speech, language, or cognitive impairment and passed a hearing screening at 25dB HL (500 Hz, 1 KHz, 2 KHz, and 4 KHz; Welch-Allyn Audioscope 3, calibrated 1/9/07; testing conducted between March & June, 2007). Two Polish and one English participant were left-handed; all others were right-handed.

Stimuli

Natural speech stimuli consisting of nonsense words (henceforth termed "words") produced by a male bilingual Polish-English speaker were recorded with a Shure SM48 microphone, Earthworks microphone, Preamp lab 101 using the software program, Sound Forge, version 4.5 (Build 49) on a Dell Dimension XPS B800 Pentium III computer with Sound Card, Creative SB Live! Basic (WDM) (sampling rate - 22,050; 16-bit resolution; channels mono). Selected nonsense word stimuli were copied to separate files and preceding and following silence intervals were removed. Selected stimuli were normalized to -20 dBVU and DC offset was set to zero (Sound Forge version 4.5).

Stimuli were potential real words in Polish and English with the exception of all nonsense words that began with /pt/, an illegal phonotactic form in English. Phoneme sequences that followed the /pt/, /pət/, /st/ and /sət/ onset segments in the words varied and words in the *pt* and *st* conditions were matched for rhyme as closely as possible.

Also, words within the pairs were matched for f_0 contour. The mean and range for the word lengths for each of the four word forms follow: three-syllable *pt*: mean = 552 ms, range 481-675 ms; two-syllable *pt*: mean = 490 ms, range 417-604 ms; three-syllable *st:* mean = 698 ms, range 633-801 ms; two-syllable *st:* mean = 622 ms, range 551-698 ms.

E-Prime software, version 1.1 (on Dell Precision PWS 670 Intel (R) computer) was used for stimulus presentation and button-press response recording. Stimuli were delivered in a sound-treated, electrically shielded room and presented through speakers. Stimuli were presented at intensity levels ranging from 51 to 63 dB SPL (Sound level meter: B&K 2203 with a B&K 4144 mic,"A" weighting).

Four *pt* conditions and four *st* conditions consisted of 100 word pairs each (250 ms inter-stimulus interval; 2000 ms inter-trial interval). Word pairs were presented in 10 blocks with 80 pairs in each block. Stimulus pairs were randomized within blocks and the blocks were presented in a random order.

EEG recordings were measured using a 65 channel Geodesic Sensor Net (Electrical Geodesic, Inc) with silver/silver-chloride (Ag/AgCL) plated electrodes. Impedance levels for each electrode were at or below 40 kOhms, an acceptable level for the high-impedance amplifiers used (Ferree, Luu, Russell & Tucker, 2001). Electrodes placed above and below the eyes monitored eye movements and eye blinks. The EEG sampling rate was 250 Hz and EEG was bandpass filtered between 0.1 and 30 Hz and referenced to Cz. The continuous recording was segmented into epochs consisting of a 200 ms baseline and a 1600 ms post-onset segment. Artifact rejection was set at +/- 70 microvolts. Data from bad channels was replaced by spline interpolation. Data was baseline corrected from -100 ms to 0 ms. and re-referenced to an average reference (Electrical Geodesic, Inc). Epochs were time-locked to the first word in the word pair.

GFP reduction was used in the first analysis to determine whether there were group and stimulus differences in 24 ms time-intervals between 64 ms and 472 ms. GFP calculates variance across electrode sites (i.e., the standard deviation of voltage from 64 channels at each data point) providing an objective reference-free measure of the time intervals of maximum brain activity (Lehmann & Skrandies, 1980; Shafer, Ponton, Datta, Morr & Schwartz, 2007; Sussman, Steinschneider, Gumenyuk, Grushko & Lawson, 2008). Current Source Density (CSD) transformations at fronto-central, lateral temporal sites and lateral frontal sites, that corresponded to time-intervals having peak GFP, were examined. The CSD data were created using the second spatial derivative of the voltage potential (Perrin et al., 1989; Sussman et. al., 2008) performed with Brain Electric Source Analysis (BESA EEG V5.1; 2005). CSD emphasizes shallow depth current flow suggesting sources in neocortex (Pernier et al., 1988; Steinschneider & Dunn, 2002; Steinschneider et al., 1992; Martin et al., 2003). Greenhouse-Geisser correction was applied for three or more factors of site and Tukey HSD was used for post hoc comparison. Only main effects of condition and language group and significant interactions involving condition will be reported.

The average amplitudes of 24 ms time-windows were used in the analysis to further reduce the data. The fronto-central P1N1P2 response was obtained from four Geodesic sites: fronto-central (site 4), left fronto-central (site 5), vertex (Cz) and right fronto-central (site 55) (Electrical Geodesic Inc.). The T-Complex was obtained from four sites: left-anterior temporal (AL, site 20), left-posterior temporal (PL, site 24), right-anterior temporal (AR, site 56), right-posterior temporal (PR, site 52). The lateral frontal waveform response was obtained from the left-lateral frontal (site 15) and right-lateral frontal (site 61) sites (See Figure 2.1 for a map of Geodesic Inc. electrode sites).

Each of the four word forms within the experimental stimuli occurred as the first word in the pair within two experimental conditions. For example, 100 words that have a /pt/ onset (e.g., /ptimɑ/) were presented as the first word in a same pair and as the first word in a different pair (see chapter 2 for greater detail). Therefore, each participant’s response to the word forms was averaged across the two experimental conditions. Hence, statistical analysis for each participant was computed on a maximum of 200 experimental trials. Figure 3.1 depicts an overlay of auditory word signatures (fronto-central site) obtained from two separate experimental conditions, which contained the same 100 word stimuli, demonstrating consistency of response.

RESULTS

TWO AND THREE-SYLLABLE *PT* WORD FORMS

***GFP pt 2 and 3-syllable words**:* Figure 3.2 shows apparent differences in GFP between 50 and 400 ms for 2 and 3-syllable pt words. A condition (2) x time (9) x group ANOVA on 24 ms time intervals between 184 and 376 revealed that the 2 and 3-syllable *pt* words differed for the intervals 304 and 328 ms. The language group response was not different.

A condition (3-syllable pt words, 2-syllable pt words) x time (5) x group analysis of GFP for the time intervals between 64ms and 160 ms revealed no main effect of group or significant interactions. Analysis of the time intervals between 184 and 376 ms revealed no main effect of condition and a significant interaction of condition x time (F (8, 176) = 2.698; p =.008; partial eta squared = .109). Post hoc comparisons of the condition x time interaction found the conditions to differ for time intervals, 304 and 328 ms (p < .01).

GFP peaks corresponded to activity in CSD maps from fronto-central, lateral temporal and lateral frontal sites.

***CSD analysis, P1N1P2 fronto-central sites, pt 2 and 3-syllable words**:* The fronto-central response to the *pt* 2 and 3-syllable words differed roughly between 200 and 350 ms. No interaction of condition x group was found suggesting that the language groups response to the 2 and 3-syllable *pt* words did not differ as shown in Figure 3.3. The Polish language group demonstrated a more negative waveform for the *pt* stimuli than the English group.

CSD analysis of the time bins between 88 ms and 256 ms in a condition (2) x time (8) x site (4) x group ANOVA of CSD data revealed no main effects and a condition x time interaction (F (7, 154) = 12.225, p = .000, partial eta squared = .357). Post hoc comparison of the condition x time interaction found the conditions to differ for the time intervals 208 (p = < .05), 232 ms (p = < .01) and 256 ms (p = <.01) as shown in Figure 3.3. Condition and language group did not interact (p = .207). Analysis of the Polish group separately in a condition (2) x time (8) x site (4) ANOVA revealed no significant main effects and a significant interaction of condition x time (F (7, 77) = 12.225, p = .011, partial eta squared = .274). Post hoc comparisons did not find conditions to differ significantly for the 8 time intervals. Analysis of the English group separately in a condition (2) x time (8) x site (4) ANOVA revealed no significant main effects and a significant interaction of condition x time (F (7, 77) = 10.473, p = .000, partial eta squared = .488). Post hoc comparisons found the conditions to differ for the final 2 time intervals 232 and 256 ms (p < .01).

CSD analysis of the time intervals between 280 ms and 496 ms in a condition (2) x time (10) x site (4) x group ANOVA found a main effect of group (F (1,22) = 9.387; p = .006; partial eta squared = .299) and a significant interaction of condition x time (F (9, 198) = 8.262; p = .000; partial eta squared = .273). Post hoc comparisons revealed 2 and 3-syllable words to differ for the time intervals 304 and 328 ms (p = < .01). Language groups differed by a mean of .09 uV (Polish mean = -.15 mV, std. err .02; English mean = -.06, std. err = .02), however, a significant interaction of condition x group was not found (p =.27) suggesting the language group response was not different for the 2 and 3-syllable *pt* words.

Each language group was analyzed separately for the time intervals between 280 and 496 ms in a condition (2) x time (10) x site (4) ANOVA follows. Analysis of the English group revealed no main effect but a significant interaction of condition x time (F (9, 99) = 5.021; p = .004; partial eta squared = .313). Post hoc comparisons found the 2 and 3-syllable words to differ significantly for the 328 ms time interval (p = .031). Analysis of the Polish group revealed no main effect but a significant interaction of condition x time (F (9, 99) = 4.269; p = .033; partial eta squared = .280). Post hoc comparisons of the condition x time interaction revealed the condition to differ for the 304 (p = .013) and 328 ms time intervals (p = .000). The auditory word signatures for the 2 and 3-syllable *pt* words did not differ for the language groups.

CSD analysis T-Complex, pt 2 and 3-syllable words: Analysis of the time intervals between 40 ms and 256 ms found the *pt* 2 and 3-syllable words to differ only for the right anterior temporal site and not the left anterior temporal sites suggesting an hemisphere effect for the *pt* word forms that did not differ for the language groups. Analysis of the posterior temporal sites between 40 and 256 ms revealed the language groups to respond differently to the 2 and 3-syllable word forms. Post hoc analysis did not find the word forms to differ significantly for either group.

Specifically, the analysis between 40 and 256 ms in a 2 (2-syllable, 3-syllable) x 10 (Time) x 4 (AL, PL, AR, PR) x group ANOVA revealed no main effect of condition or group and a significant interactions of condition x time x site (F (27, 594) = 2.424; p = .015; partial eta squared = .099). A step-down procedure analyzed anterior and posterior sites separately. Analysis of the left and right-posterior temporal sites in a condition (2) x time (10) x sites (2) x group ANOVA revealed no main effects of condition or group and a significant condition x group interaction (F (1, 22) = 8.235; p = .009; partial eta squared = .272) with the Polish group having a more negative response to the 2-syllable *pt* words as shown in Figure 3.4. Post hoc analysis, however, did not find conditions to differ significantly by group. Figure 3.5 illustrates a large negativity at 200 ms from the left posterior temporal site in response to the *pt* words that differs for the language groups. Each language group analyzed separately in a condition (2) x time (10) x sites (left, right-posterior temporal sites) ANOVA revealed no significant main effect or interactions. Analysis of the left and right-anterior temporal sites in a condition (2) x time (10) x site (2) x group ANOVA revealed no main effect of group or condition and a significant interaction of condition x site (F (1, 22) = 7.307; p = .013; partial eta squared = .249). A significant interaction of condition x time x site was also found (F (9, 198) = 4.467; p = .000; partial eta squared = .169). Post hoc analysis found the conditions were not significantly different at either site. An analysis of the left-anterior temporal site in a condition (2) x time (10) x group ANOVA followed the three-way interaction and did not find significant main effects or interactions. The same analysis of the right-anterior temporal site revealed no main effect of group, a main effect of condition (F (1, 22) = 8.937; p = .007; partial eta squared = .289) and a significant interaction of condition x time (F (9, 198) = 2.75; p = .005; partial eta squared = .111). Conditions were significantly different for the time intervals 184, 232 and 256 ms (p < .01). In sum, the *pt* conditions showed differences between 40 and 256 ms from right, but not left anterior temporal sites as shown in Figures 3.6 through 3.8.

Analysis of the later time window between 280 and 496 ms did not find the 2 and 3-syllable *pt* words to differ significantly for any of the 10 time intervals and language groups did not differ.

A condition (2) x time (10) x site (4) x group ANOVA within the time window between 280 and 496 ms revealed no main effect of group or condition and no significant interactions. A separate analysis of anterior and posterior temporal sites follows to explore hemisphere. Analysis of the posterior temporal sites in a condition (2) x time (10) x site (2) x group ANOVA revealed no main effects, a significant interaction of condition x time (F (19, 198) = 3.861; p = .000; partial eta squared = .149 and a significant interaction of condition x time x group (F (19, 198) = 1.932; p = .049; partial eta squared = .081. An analysis of the left and right posterior sites, separately, follows the three-way interaction. Analysis of the left posterior temporal site in a condition (2) x time (10) x group ANOVA revealed no main effects and a marginally significant interaction of condition x time (F (19, 198) = 1.913; p = .052; partial eta squared = .080. Analysis of the right posterior temporal site revealed no main effects and a significant interaction of condition x time (F (19, 198) = 3.571; p = .000; partial eta squared = .140. Post hoc comparisons found that the conditions did not differ significantly for the 10 time intervals. Analysis of the anterior temporal sites in a condition (2) x time (10) x site (2) x group ANOVA revealed no main effects and a significant interaction of condition x sites (F (1, 22) = 5.572; p = .028; partial eta squared = .202. Post hoc comparisons revealed that the 3-syllable condition differed significantly for the temporal sites (p < .01). Figure 3.6 depicts a different pattern for the right and left sites with a more positive response to the 3-syllable *pt* words compared to the 2-syllable at the right anterior site and a reverse pattern at the left anterior site. Condition, time and group did not interact (p = .9)

CSD analysis lateral frontal, pt 2 and 3-syllable words: A different pattern of responding to the *pt* 2 and 3-syllable words by the English and Polish participants was found from the right-lateral frontal site. Conditions differed significantly for the time intervals 208 and 256 ms for the English participant group, but not the Polish group from the right-lateral frontal site. Figure 3.9 illustrates a greater variability of response to the 2-syllable words for the English participants compared to the Polish participants.

Analysis of the time intervals between 40 ms and 256 ms in a condition (2) x time (10) x site (left-lateral frontal, right-lateral frontal) x group ANOVA revealed no main effect of group or condition and a significant interactions of condition x time x site (F (9,198) = 2.551; p = .009; partial eta squared = .104) and condition x time x site x group (F (9, 198) = 2.98; p = .002; partial eta squared = .119). A step down analysis of the left lateral frontal site revealed no main effect or interactions. Analysis of the right-lateral frontal site revealed marginal effect of language group (F (1, 22) = 4.177; p = .053; partial eta squared = .16), no main effect of condition and a significant interaction of condition x time (F (9,198) = 6.099; p = .000; partial eta squared = .217) and condition x time x group (F (9,198) = 1.987; p = .043; partial eta squared = .083). Post hoc analysis of the condition x time interaction found conditions to differ significantly at time intervals 208 ms (p < .05) and 256 ms (p < .01). The language groups' data were examined separately from the right lateral frontal site to investigate the three-way interaction condition x time x group. Analysis of the Polish group (2 x 10 ANOVA) revealed no main effect of condition, main effect of group or a significant interaction. In contrast, analysis of the English group at the right-lateral frontal site revealed a significant condition x time interaction ($F (9, 99) = 6.11$; $p = .000$; partial eta squared = .357). Post hoc comparisons of the condition x time interaction for the English group found the 2 and 3-syllable *pt* words to differ for the time intervals 208 and 256 ms ($p < .05$). As depicted in Figure 3.9, greater variability in response to the 2-syllable *pt* words for the time window between 208 and 256 ms is evident for the English participants compared to the Polish participants.

Analysis of the time intervals between 280 and 496 in a 2 (Condition) x 10 (Time) x 2 (left lateral frontal, right lateral frontal) x group revealed no main effect of language group or condition and no significant interactions.

Summary of ERP response to 2 and 3-syllable pt words:

GFP analysis showed the 2 and 3-syllable *pt* word forms to diverge during the time intervals 304 and 328 ms and the response was the same for both language groups.

CSD analysis of the midline fronto-central P1N1P2 response to the *pt* 2 and 3-syllable words revealed that the words differed roughly between 200 and 350 ms. Furthermore, the Polish group demonstrated a more negative waveform for the *pt* stimuli, however, the P1N1P2 complex in response to the 2 and 3-syllable *pt* words did not differ for the language groups. Therefore, native-English and native-Polish listeners responded to 2 and 3-syllable *pt* words within the fronto-central P1N1P2 complex in the same manner.

CSD analysis of the T-Complex response revealed a hemisphere effect in that a greater distinction in response to the 2 and 3-syllable *pt* words was found at the right anterior temporal site compared to the left anterior temporal site during the early time window 40 to 256 ms. Analysis for the later time window between 280 and 496 ms found a different response to the 3-syllable *pt* words at the left and right anterior temporal sites. Also, native-language pattern of speech perception were reflected within the T-Complex because the language groups responded differently to the 2 and 3-syllable *pt* words between 40 and 256 ms from the right and left posterior temporal sites. Analysis of the later time window between 280 and 496 ms did not find the language groups to respond differently to the 2 and 3-syllable *pt* words.

A language group effect was found from the lateral frontal sites because the English participants' response to the 2 and 3-syllable *pt* words differed for the time intervals 208 and 256 ms from the right-lateral frontal site but the Polish participants response did not. A greater variability of response was evident for the English listeners to the *pt* 2-syllable words between 208 and 256 ms relative to the Polish listeners suggesting more efficient processing for the 2-syllable stimuli.

TWO AND THREE-SYLLABLE *ST* WORD FORMS

***GFP st 2 and 3-syllable words**:* Figure 3.10 shows apparent differences in GFP between 50 and 500 ms for 2 and 3-syllable *st* words, although they appear to be somewhat larger for the Polish group than the English group. A condition (2) x time (7) x group (2) ANOVA on 24 ms time intervals between 328 and 472 revealed a significant interaction of condition and time. Post hoc comparisons, however, did not find the seven time intervals to differ significantly. No other main effects or significant interactions were found for the time window analyzed between 64 and 160 ms and 184 and 256 ms.

GFP peaks in response to the 2 and 3-syllable *st* word forms were analyzed for time bins between 64ms and 160 ms in a condition (2) x time (5) x group ANOVA and between 184 and 256 ms in a condition (2) x time (4) x group ANOVA revealing no main effect of condition or group and no significant interactions. Analysis of the GFP between 328 to 472 ms found no main effects, but a significant interaction of condition x time (2 x 7 x group ANOVA; F (6, 132) = 3.390; p =.004; partial eta squared = 134). Post hoc analysis of the condition x time interaction did not find conditions to differ for the seven time intervals.

GFP peaks in response to the *st* stimuli corresponded to activity in CSD maps from fronto-central, lateral temporal and lateral frontal regions.

CSD analysis, P1N1P2 fronto-central, st 2 and 3-syllable words: Figure 3.11 shows that the auditory signature waveform differs for the 2 and 3-syllable *st* words particularly between 200 and 500 ms.

Analysis of time bins between 88 ms and 256 ms in a condition (3 syllable, 2 syllable) x time (8) x site (4) x group ANOVA revealed no main effect of condition or group, but a significant interaction of condition x group (F (1,22) = 9.86, p = .005, partial eta squared = .309) and condition x time (F (7, 154) = 8.051, p = .000, partial eta squared = .268). Post hoc comparisons of the condition x group interaction revealed the groups to be significantly different for the 2-syllable *st* word forms, only (p = .017) as shown in Figure 3.12 and 3.13. Also, post hoc comparisons of the condition x time interaction found conditions to be significantly different for the time intervals between 232 ms and 256 ms (p = < .01). Analysis of the groups separately in a condition (2) x time (8) x site (4) ANOVA revealed no significant main effect or interactions for the Polish group. Analysis of the English group revealed a significant main effect of condition (F (1,11) = 10.309, p = .008, partial eta squared = .484) and a significant interaction of condition x time (F (7, 77) = 7.424, p = .000, partial eta squared = .403). Post hoc comparisons of the condition x time interaction for the English group found the s*t* 2 and 3-syllable words to differ significantly for the five time intervals between 160 and 256 ms (p < .01). The main effect of condition for the English group found the 2 and 3-syllable words to differ by a mean of -.04 µV (mean 3-syllable = -.014, std err. = 013; mean 2-syllable = .026, std err. = .013).

Analysis of the intervals between 280 and 496 in a condition (2) x time (10) x site (4) x group ANOVA, revealed a main effect of group (F (1, 22) = 12.564; p = .002; partial eta squared = .364), a significant interaction of condition x group (F (1,22) = 5.852; p = .024; partial eta squared = .210), a significant interaction of condition x time (F (9,198) = 15.347; p = .0; partial eta squared = .411) and a significant interaction of condition x site x group (F (3, 66) = 3.320; p = .037; partial eta squared .131). Post hoc comparisons of the condition x group interaction revealed the groups to be different for the 2-syllable *st* forms, only (p < .01) and comparisons of the condition x time interaction revealed the conditions to differ significantly for the time intervals 328 (p = .000), 352 (p = .000), 400 (p = .023), 424 (p = .000), and 448 ms (p=.015). The three-way interaction of condition x site x group was followed by an analysis of site. Analysis of the midline fronto-central site found a main effect of group (F (1, 22) = 9.243; p = .006; partial eta squared = .296) and a significant interaction of condition x time (F (9, 198) = 11.750; p = .000; partial eta squared = .348) and condition x group (F (1, 22) = 9.068; p = .006; partial eta squared = .292). Post hoc comparisons of the condition x group interaction revealed the groups to be different for the 2-syllable s*t* forms (p < .01) as shown in Figure 3.12. Post hoc comparisons of the condition x time interaction revealed the conditions to differ significantly for the time intervals 328 (p = .000), 352 (p = .000), 424 (p = .000), and 448 ms (p=.03). Analysis at the left fronto-central site found a main effect of group (F (1, 22) = 10.696; p = .003; partial eta squared = .327) and a significant condition x time interaction (F (9,198) = 10.701; p = .000; partial eta squared = .327). Post hoc comparison of the condition x time interaction revealed the conditions to differ significantly for the time intervals 280 (p =. 034), 328 (p = .008), 352 (p = .000), 400 (p = .045) and 424 (p = .001). Analysis of the response at Cz found no main effect of group or condition and a significant interaction of condition x time (F (9, 198) = 7.411; p = .000; partial eta squared = .252). Post hoc comparisons of the condition x time interaction revealed the conditions to differ significantly for the time intervals 424 (p < .01) 448 ms (p=.042). Analysis of right fronto-central site found no main effect of group or condition, but a significant interaction of condition x time (F (9,198) = 9.311; p = .000; partial eta squared = .297). Post hoc comparisons of the condition x time interaction revealed the conditions to differ significantly for the intervals 328 (p < .01), 352 (p < .01) and 424 (p < .05).

Separate analysis of site revealed that the language groups responded differently to the conditions at only the midline fronto-central site. Furthermore, the conditions differed for a longer time window at midline fronto-central and left fronto-central sites than right and vertex sites.

Separate analysis of the groups for the intervals between 280 and 496 in a condition (2) x time (10) x site (4) ANOVA follows. Analysis of the English group revealed a main effect of condition (F (1, 11) = 5.321; p = .042; partial eta squared = .326) and a significant interaction of condition x time (F (9, 99) = 8.628; p = .000; partial eta squared = .440). Post hoc comparisons of the condition x time interaction revealed the *st* 2 and 3-syllable conditions to differ for the four time intervals, 280, 400, 424 and 448 ms (p < .01). The main effect found the conditions to differ by -.033 µV (3-syllable mean = -.068, std. err. = .012; 2-syllable mean = -.035, std. err. = .017). Analysis of the Polish group revealed no main effects, but a significant interaction of condition x time (F (9, 99) = 7.804; p = .000; partial eta squared = .415). Post hoc comparisons found the conditions to differ for the time intervals 376 and 400 ms (p < .01).

The separate language group analyses using all four sites for the time window between 280 and 496 ms revealed that each language groups' response reflected detection of the distinction between the 2 and 3-syllable *st* words and there was an overlap for the language groups at the time interval, 400 ms, where detection of the distinction reached significance.

CSD Analysis, T-Complex, st 2 and 3-syllable words: Analysis of the T-Complex response to the 2 and 3-syllable word forms for the time-window 40 through 256 revealed no main effects or significant interactions (conditions (2) x time (10) x site (4) x group ANOVA).

Analysis of the time-window between 280 and 496 in a condition (2) x time (10) x site (AL, AR, PL, PR) x group ANOVA revealed no main effects and a significant interaction of condition x time (F (9, 198) = 2.695; p = .05; partial eta squared = .109) and condition x time x site (F (27, 594) = 2.978; p = .005; partial eta squared = .119). Conditions did not differ significantly for any of the 10 time intervals on post hoc comparisons. An analysis of the anterior and posterior temporal sites follows the three-way interaction. Analysis of the posterior temporal sites and the anterior temporal sites separately revealed no main effects and a significant interaction of condition x time (Posterior temporal: 2 x 10 x 2 x group F (9, 198) = 3.001; p = .002; partial eta squared = .120; Anterior temporal: 2 x 10 x 2 x group, (F (9, 198) = 4.499; p = .0; partial eta squared = .17). The conditions did not differ significantly for any of the 10 time intervals on post hoc comparisons.

CSD Analysis, Lateral Frontal, st 2 and 3-syllable words: Analysis of the right and left lateral frontal sites (2 x 10 x 2 x group ANOVA) between the time intervals of 40 and 256 ms revealed no main effects or significant interactions.

Analysis of the later time window between 280 and 496 ms revealed no main effect and a significant interaction of condition x time (F (9, 198) = 10.056; p = .000; partial eta squared = .314). Post hoc comparisons revealed the conditions to differ significantly for the time intervals 400, 424 and 448 ms (p < .01).

Summary of ERP response to 2 and 3-syllable st words:

GFP differences in response to the 2 and 3-syllable *st* stimuli were not significant.

CSD analysis found that the P1N1P2 response differed for the 2 and 3-syllable *st* words particularly between 200 and 500 ms. While both language groups distinguished the 2 and 3-syllable stimuli, the language groups responded differently to the 2-syllable s*t* words for the time-window between 88 and 496 ms.

CSD analysis of the T-Complex response to the 2 and 3-syllable *st* words did not find the conditions to differ significantly for any of the time intervals and the language group response was not different. CSD analysis of the T-Complex response to the *pt* and *st* word forms was more reflective of the native-language pattern of speech perception in that the language groups differed for the *pt* 2 and 3-syllable words (native-language contrast in the Polish language, only) but did not differ for the *st* 2 and 3-syllable words (native-language contrast in Polish and English).

CSD analysis of the lateral frontal sites revealed the *st* 2 and 3-syllable words to differ significantly for the time intervals between 400 and 448 ms.

DISCUSSION

P1N1P2 Complex to non-native patterns

The most important finding in the current research was that sensitivity to native-language phonotactic patterns was not reflected in the auditory word signatures of the P1N1P2 complex. The auditory word signatures, which distinguish the 2 and 3-syllable *pt* words contained the word onset /pt - pət/ contrast that is a linguistic contrast in the Polish language, but not the English language. Nevertheless, the auditory word signatures for the 2 and 3-syllable *pt* words were highly similar for the native-Polish and native-English listeners. As reported in Chapter 2, English participants in a behavioral task were unable to distinguish words beginning with /pt/ or /pət/ within a syllable identification task and this inability was reflected in a late positive ERP component (i.e., LPC) peaking at approximately 500 ms. These same English participants demonstrated auditory word signature responses (P1N1P2 complex) to the 2 and 3-syllable *pt* words that were not different from the Polish participants. These results suggest that English listeners (as well as the Polish listeners) have access to the acoustic features distinguishing the word onset non-native contrast, /pt/ versus /pət/.

Previous studies suggest that the P2 obligatory auditory evoked potential (AEP) is sensitive to experience. A group of skilled musicians demonstrated a greater P2 amplitude in response to musical tones relative to non-musicians (Shahin et al., 2003). Also, a short period of auditory exposure and auditory discrimination training resulted in modified P2 amplitude/latency responses (Tremblay & Kraus, 2002; Sheenan et al., 2005; Reinke, 2003). Exposure to one’s native-language provides the ultimate practice of a skill, speech perception of native-language contrasts. Why, then, does the Polish group not demonstrate a greater P2 response to the 2-syllable /pt/ words? Polish listeners have experience with the contextual acoustic features of word onset /pt/ during development, whereas English listeners do not. Consider that musicians' life-long experience with musical tones modified P2 amplitudes, but Polish listeners' native-language experience with contextual features of the /pt/ onset cluster did not affect the P2 amplitudes of the P1N1P2 complex relative to the English listeners. During development, both musicians and non-musicians are exposed to musical tones of the culture but one group has greater experience than the other. In contrast, phonotactic constraints of English prevent word onset /pt/ and, hence, the acoustic features of the word onset cluster do not occur in the language. Hence, the greater P2 response by the musicians in Shahin et al., may reflect greater attention to the musical tones. A stimulus must be salient to be targeted for attention (Coull, 1998). We suggest that exposure to the acoustic features of the word onset /pt/ cluster within the language during development is necessary for the features of the cluster to be salient. Therefore, detection of the acoustic features of the /pt/ onset cluster is reflected within the P1N1P2 complex by both language groups, but differing degrees of attention to the stimuli is not. The English group does not demonstrate greater attention to the 2-syllable pt words because they are not salient. The Polish group does not show increased levels of attention to either the 2 or 3-syllable word. We continue this discussion below when we compare the fronto-central response to the *pt* and *st* contrasts.

Ceponiene et al., suggested that the P2 component might, partially, reflect modulation of the access to speech perception (2002, p 402). Also, Tremblay and Kraus suggested that P2 might reflect " a pre-attentive alerting mechanism that contributes to improved perception, but is not necessarily associated with acoustic or linguistic encoding" (2002, p 570). We found no evidence within our results to support the claim that P2 reflects modulation of access to speech perception or that the P2 reflects a pre-attentive alerting mechanism that contributes to improved perception because the Polish listeners' response did not show a difference in P2 despite years of experience with the *pt* contrast. Ceponiene and colleagues' model infers a serial pathway with neural generators underlying the P2 response to lead to neural processes contributing to speech perception. A parallel pathway with neural generators different than those underlying the P2 component has been suggested for neural processes contributing to speech perception (Winkler & Naatanen; 1999; Davis & Zerlin, 1966). Our results suggest a parallel pathways because language-group differences in response to the 2 and 3-syllable *pt* words were found from lateral temporal sites but not fronto-temporal sites during an early time window between 40 and 256 ms.

T-Complex to non-native patterns

The English and Polish native-language groups responded differently to the 2 and 3-syllable *pt* words between 40 and 256 ms from the posterior temporal sites (right and left). The Polish group demonstrated a more negative waveform response for the 2-syllable words relative to the 3-syllable *pt* words with the opposite pattern for the English group. A T-Complex response indexing neural processes important for speech perception has been previously demonstrated (Tonnquist-Uhlen et al., Shafer, Schwartz & Martin, in press). Furthermore, a lateral temporal effect of auditory experience is consistent with the findings of Shahin et al., (2003) and Reinke et al., (2003). Whereas primary and secondary auditory cortices on the supra-temporal plane are major contributors to the

P1N1P2 complex, secondary auditory cortex on the lateral surface of the superior temporal gyrus is a major contributor to the T-Complex. Language group differences would be expected at higher levels of speech processing within secondary auditory cortex. A different response to the 2-syllable *pt* word forms (illegal word form in English) within the T-Complex by the language groups began within an early time window between 40 to 256 ms time window. This effect of native-language at short latencies suggests that parallel processes within secondary auditory cortex are active in native-language speech perception.

The English and Polish groups also responded differently to the *pt* 2 and 3-syllable words from the right-lateral frontal site and the English participants showed large variability of response relative to the Polish participants between 200 ms and 275 ms from the right-lateral frontal site. These results also suggest a native-language effect.

The P1N1P2 response to native-language patterns

The P1N1P2 response (CSD) to the *st* word forms, which contain the word onset contrast /st - sət/ found in both English and Polish, differed for the native-language groups. The *st* 2 and 3-syllable word forms were distinguished within the P1N1P2 roughly between 230 and 475 ms for both language groups, however, a different response to the 2-syllable *st* words was found for the Polish and English listeners. The 2-syllable *st* words differed for the language groups with the Polish group showing a more negative response for the time-window between 88 ms and 496 ms. This negativity may be the processing negativity (PN) observed in previous studies when participants were instructed to attend to a stimulus (Shafer, Ponton, Datta, Morr, & Schwartz, 2007; Steinschneider and Dunn, 2002). Language group processing differences to the *st* contrast were, also, found in our target word data analysis (second word of the word pair) reported in Chapter 2. Specifically, the English and Polish participants' response demonstrated amplitude and topographical differences to the *st* contrast within a late positive component at approximately 500 ms reflecting processing differences to a contrast that occurs in both languages. Also, a large GFP response by the English group to the 2-syllable *st* target word of the different pair, not seen in the Polish participants' data, is suspected to reflect ease of processing by the English group (Chapter 2). A processing negativity within the fronto-central response to the 2-syllable *st* words greater for the Polish group in the prime word data, in addition to processing differences to the *st* contrast in the target word data, suggest that the language groups have differing linguistic experience with the 2-syllable *st* stimuli. The different response to the 2-syllable *st* words and not the 3-syllable words favors an explanation of differing linguistic experience over a differing response to our speaker's English accent (see Chapter 2 for details regarding the speaker of the experimental stimuli). Taken together, it appears the English listeners were more assured of perception of the 2-syllable *st* word forms and the Polish listeners engaged more attentional resources to the 2-syllable *st* stimuli. The reasons for these differences will have to await a comparative linguistic analysis of these phoneme sequences within words in the two languages. It will also be important to replicate this finding in another group of English and Polish participants to ensure that the finding is related to language experience differences rather than some unknown group difference. Future research may find that words containing the /st/ onset in English (i.e., stop, star, stair, step, stick) are more frequent, more common, or are learned earlier than words beginning with /st/ in Polish.

The T-Complex response to native-language patterns

The T-Complex response to the *st* contrast was not different for the native-language groups. Furthermore, the contrast within the 2 and 3-syllable *st* words was not evident within the T-Complex, whereas the 2 and 3-syllable *pt* stimuli were distinguished within the T-Complex response. The N1 amplitude being smaller in response to high frequency sounds is argued to be a result of the tonotopic organization of the auditory cortex on the superior temporal plane with neurons responding best to high frequency stimuli being deeper than neurons responding to low to mid frequency stimuli (Wunderlich and Cone-Wesson, 2001). The tonotopic organization of auditory cortex may explain our findings from lateral temporal sites.

Comparison of the P1N1P2 complex and the T-Complex response to a linguistic and non-linguistic contrast

Interestingly, the P1N1P2 response did not reflect the native-language patterns of speech perception in that the response to the *pt* stimuli was the same for the language groups and the response to the *st* stimuli, which was phonologically contrastive for both groups, was different for the language groups. The N1 and P2 of the P1N1P2 response is affected by attention to stimuli (Picton & Hillyard, Steinschneider & Dunn, 2002) and differing linguistic experience may have affected the language groups' response to the *st* stimuli. On the other hand, the English language group was perceptually unaware of the distinction within the *pt* contrast, as discussed in Chapter 2, therefore, the 2-syllable *pt* words could not be targeted for attention by our English participants (Coull, 1998).

In contrast to the fronto-central response, the T-Complex reflected the native-language patterns of speech perception with a different language group response for the *pt* but not the *st* stimuli consistent with the behavioral patterns of speech perception reported in Chapter 2. An unexpected result was the time window for the language group differences within the T-Complex response beginning at a relatively short latency. These results suggest parallel processes involved in sound representation for the perception of speech.

Hemisphere Effect

A right hemisphere effect was found for the English and Polish listeners in response to the *pt* words. A distinction was found between the 2 and 3-syllable *pt* words from the right-anterior temporal site between 180 and 275 that was not found from the left hemisphere site. A hemisphere effect was not found for the *st* contrast. These results are consistent with Zhang et al., (2005) who found left-hemisphere processing for a /ra-la/ sound contrast and not a /ba-wa/ contrast.

CSD analysis

CSD analysis was applied to the data because the analysis has been shown (animal research) to correspond to extracellular potentials associated with current sinks and sources within cortical laminae (Steinschneider & Dunn, 2002; Steinschneider et al., 1992). CSD analysis emphasizes cortical sources by removing information that probably has origins in deeper subcortical regions.

Analysis of raw voltage data, reported elsewhere, also found that the language groups' response to the 2 and 3-syllable *pt* words within the P1N1P2 response were not different. On the other hand, the results of the raw voltage data analysis and CSD analysis found different results pertaining to the language groups' response to the 2 and 3-syllable *st* words within the fronto-central response. Specifically, CSD analysis found group differences within to the 2 and 3-syllable *st* stimuli and raw voltage data analysis did not suggesting that CSD analysis reveals higher-level processing reflecting native-language patterns of speech perception and fine-grained aspects of native-language speech perception.

Conclusion

In conclusion, our findings suggest that language experience, in general, is sufficient to allow for precise registration of the acoustic-properties of speech at the level of the P1N1P2 complex. Furthermore, a parallel system of processing reflected within the T-Complex component may reflect neural networks contributing to native-language speech perception.

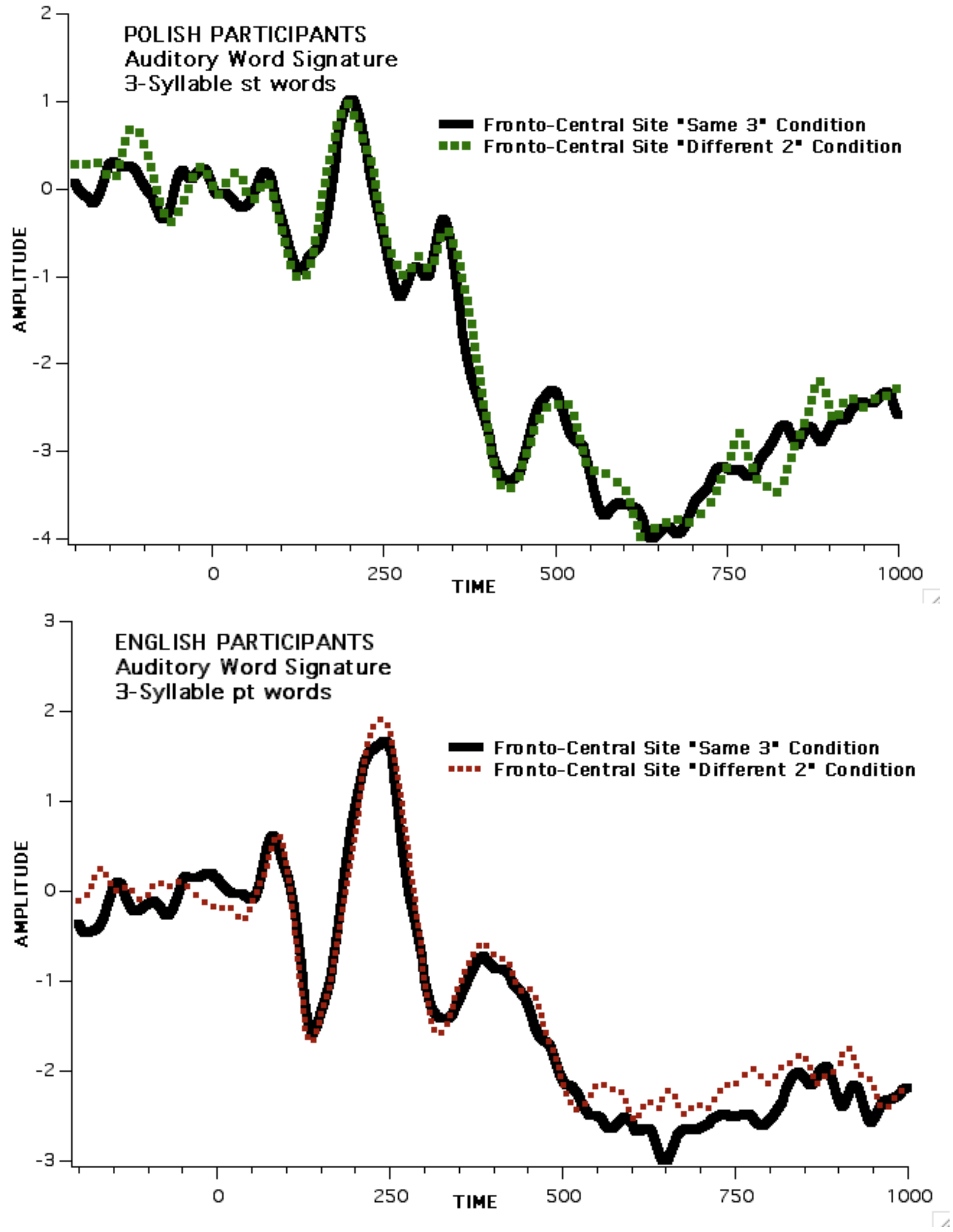


Figure 3.1 depicts the auditory word signatures for the 3-syllable *st* word forms (top) and the 3-syllable *pt* word forms (bottom) from the fronto-central site. An overlay of the auditory word signatures from two experimental conditions demonstrates the consistency of response

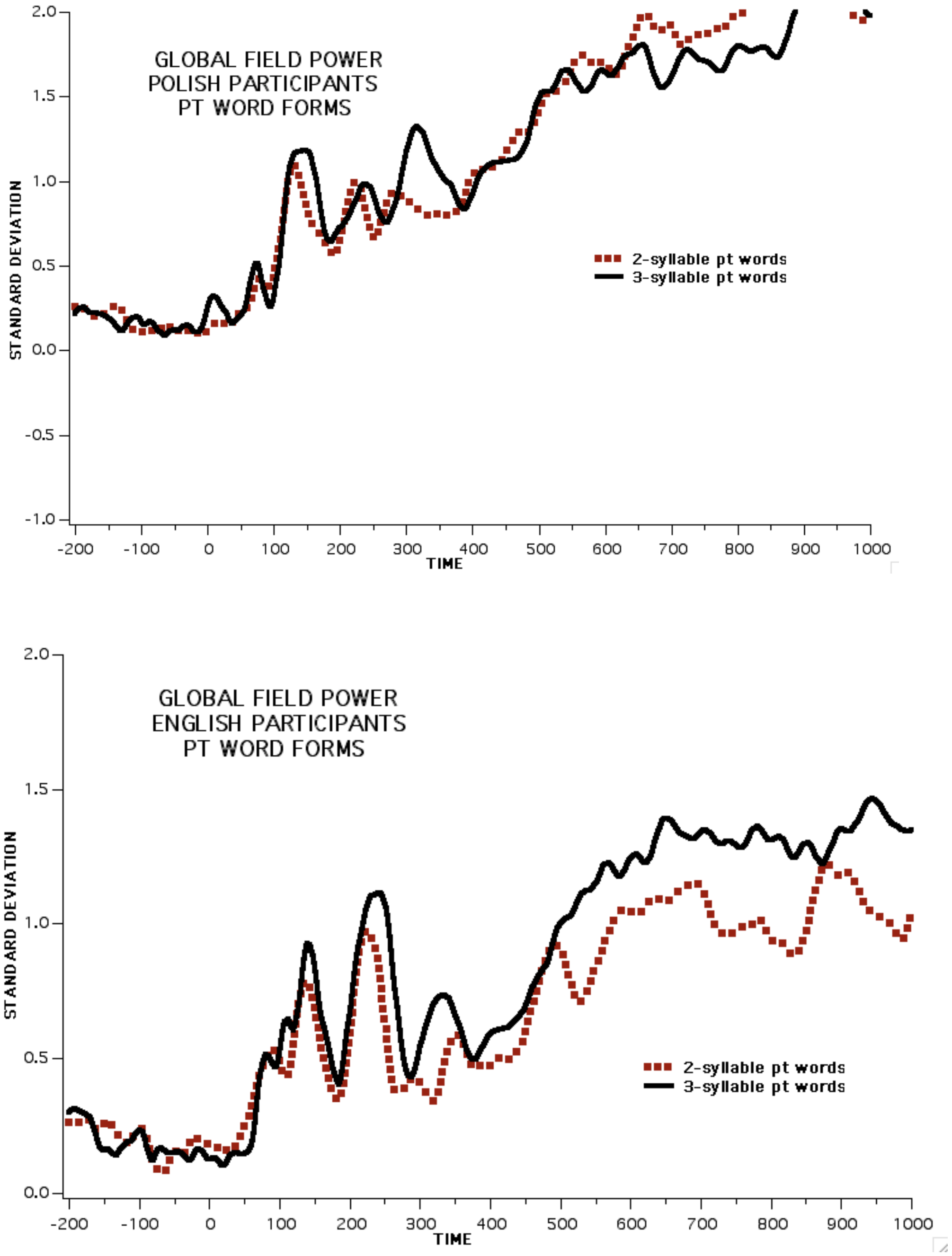


Figure 3.2 illustrates the GFP analysis of the response to the 2 and 3 syllable *pt* words for Polish (top) and English (bottom) listeners

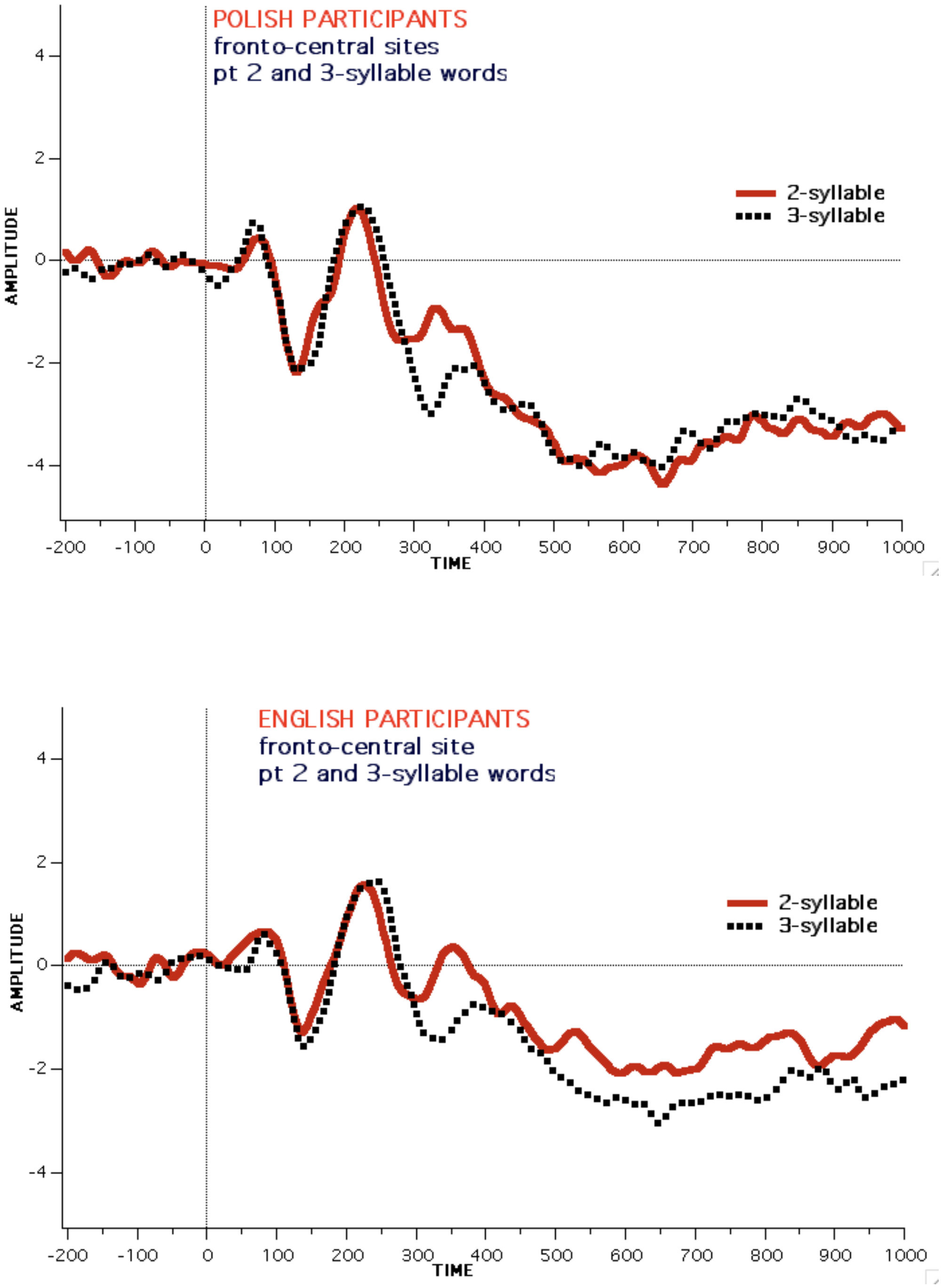


Figure 3.3 depicts the grand mean auditory waveforms in response to the 2 and 3-syllable *pt* words from the midline fronto-central site for the Polish (top) and English (bottom) participants. Notice the highly similar auditory word signature patterns for the 2 and 3-syllable words for the language groups. The raw voltage waveforms are shown

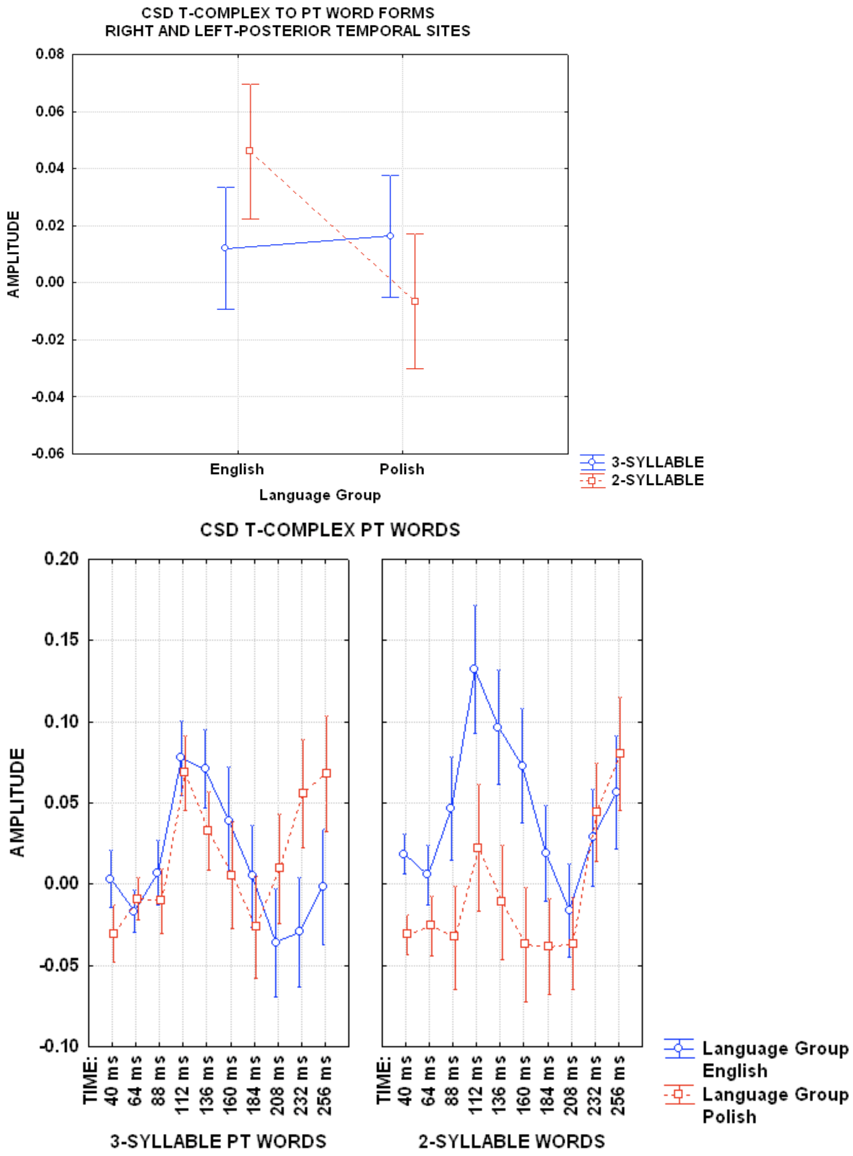


Figure 3.4 (top) illustrates a different response to the 2-syllable *pt* word forms by English and Polish groups from posterior temporal sites between 40 and 256 ms. (bottom) The same response is depicted across time

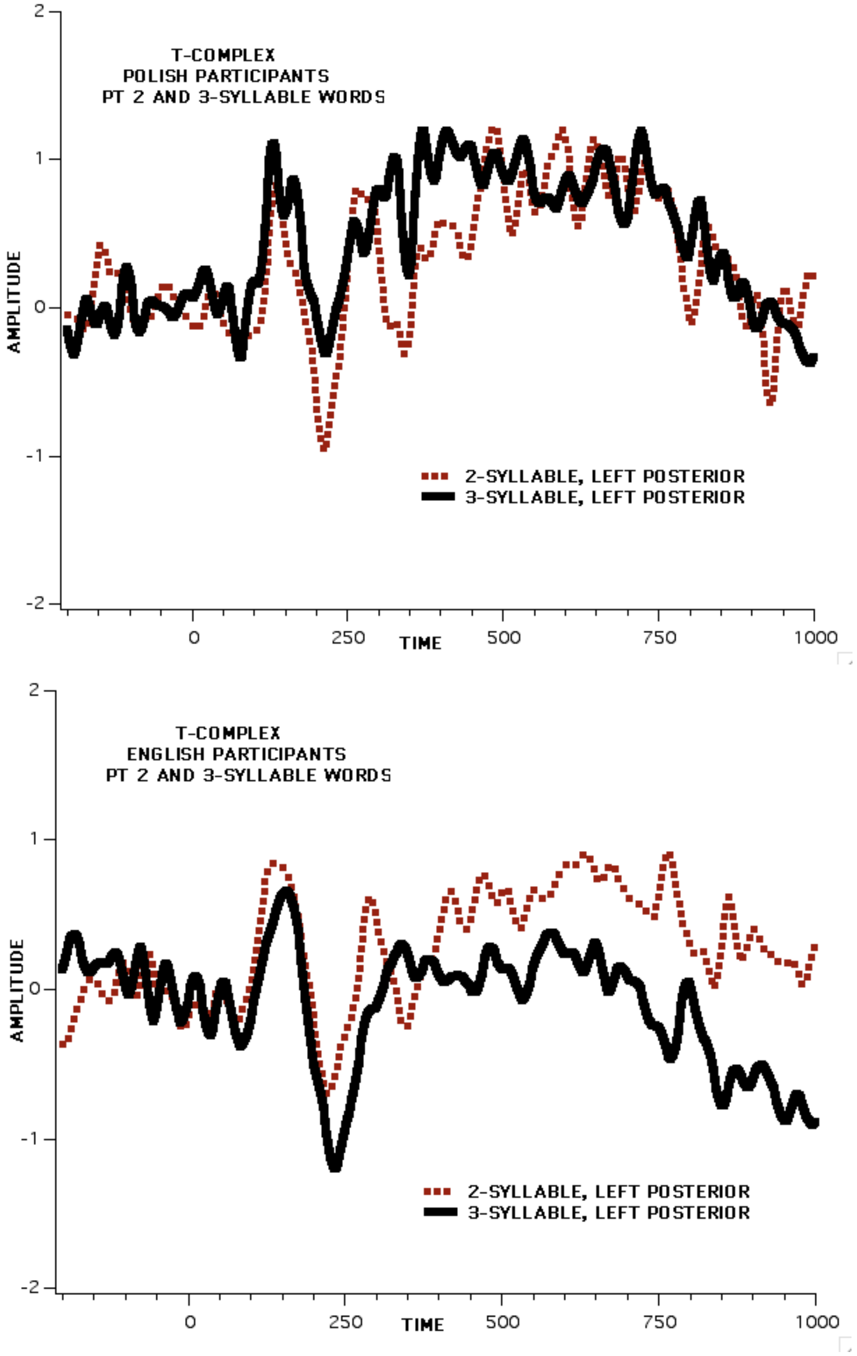


Figure 3.5 illustrates a more negative T-Complex response at 200 ms to the 2 syllable *pt* words by the Polish participants (top) but not the English participants (bottom). Raw voltage data is shown. Each waveform is the grand mean from one of the two averaged conditions

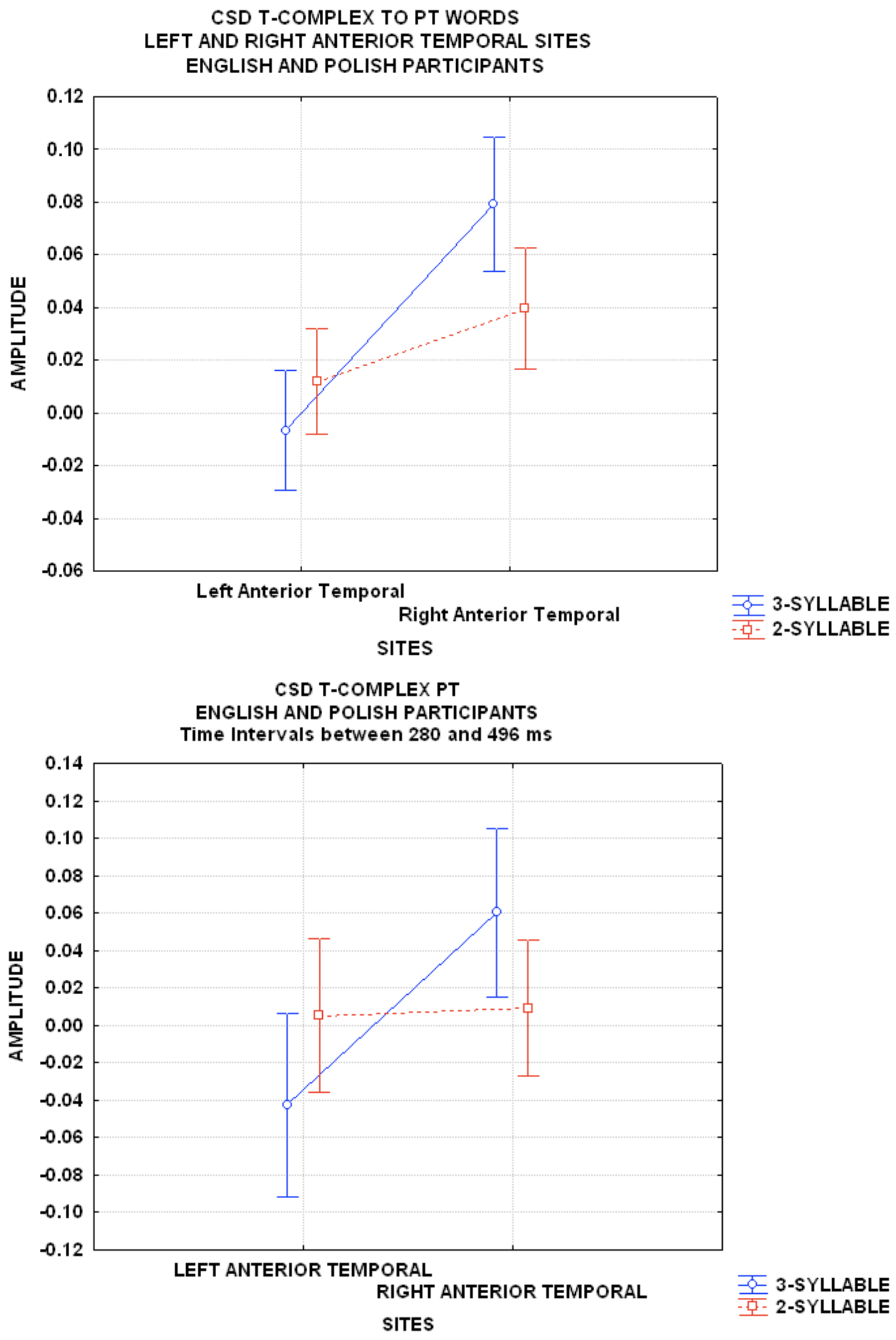


Figure 3.6 (top) depicts a right-anterior temporal hemisphere effect in response to the 2 and 3-syllable *pt* words in Polish and English participants within the time window between 40 and 256 ms. (bottom) A more positive response to the 3-syllable *pt* word than the 2-syllable word at the right anterior site is shown with a reverse pattern at the left anterior site

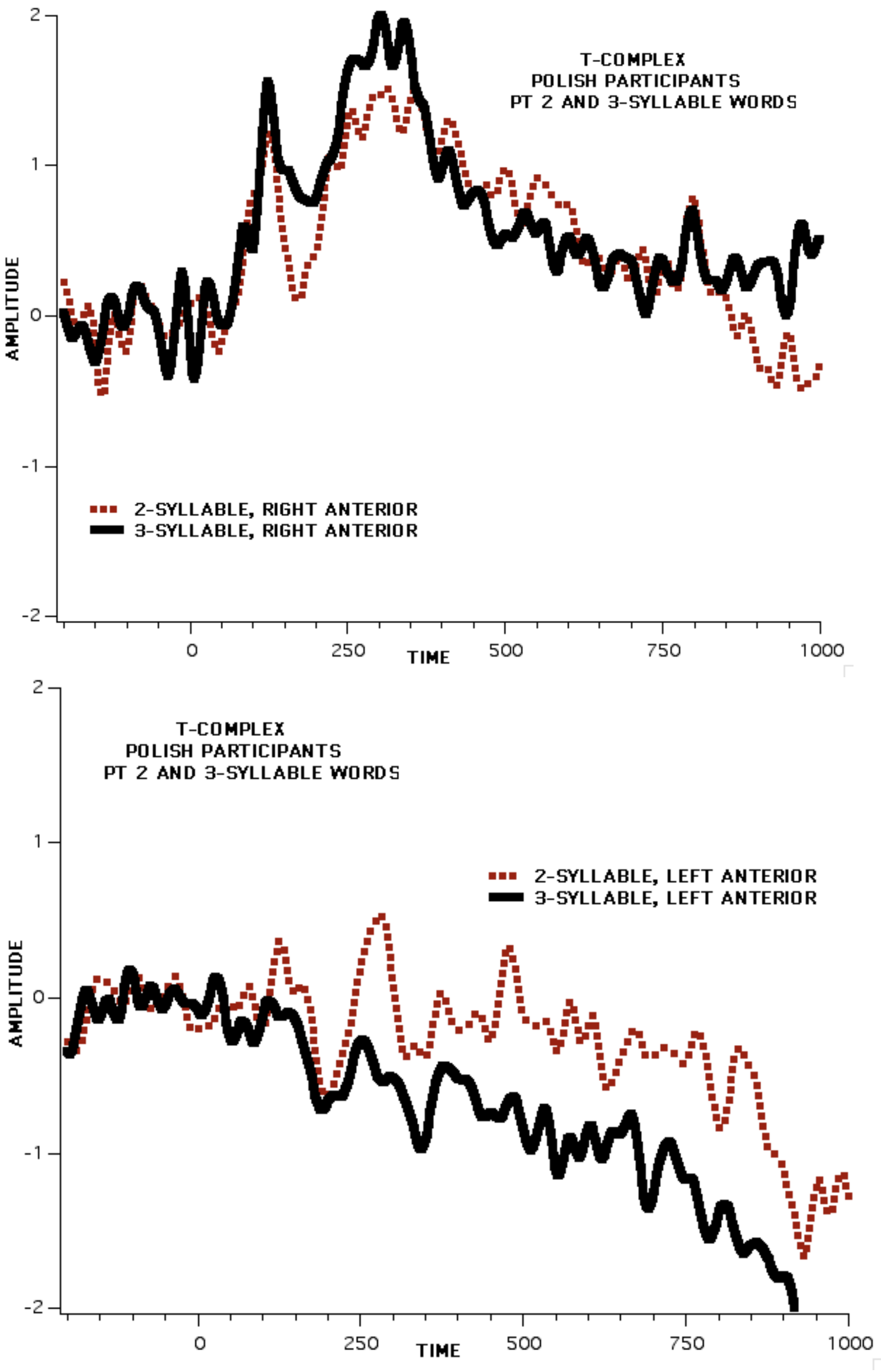


Figure 3.7 depicts a greater negative T-Complex response to the 2-syllable *pt* words at 200 ms from the right anterior temporal site (top) but not from the left anterior temporal site (bottom) for Polish participants. Raw voltage data is shown for one of the two averaged conditions

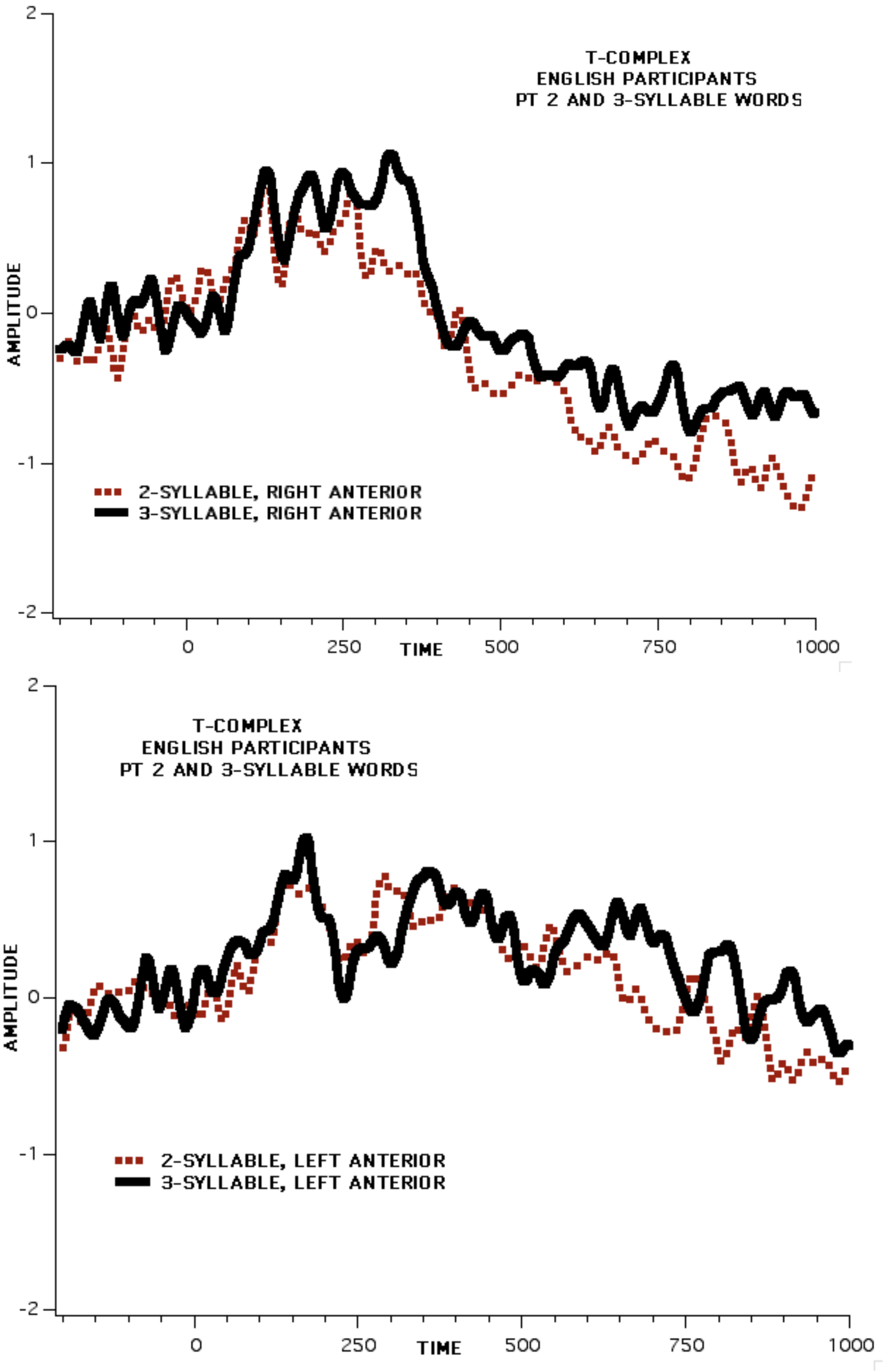


Figure 3.8 depicts a small negative T-Complex response to the 2-syllable pt words at 180 ms from the right anterior temporal site (top) for the English participants but not from the left anterior temporal site (bottom). Raw voltage data is shown for one of the two averaged conditions

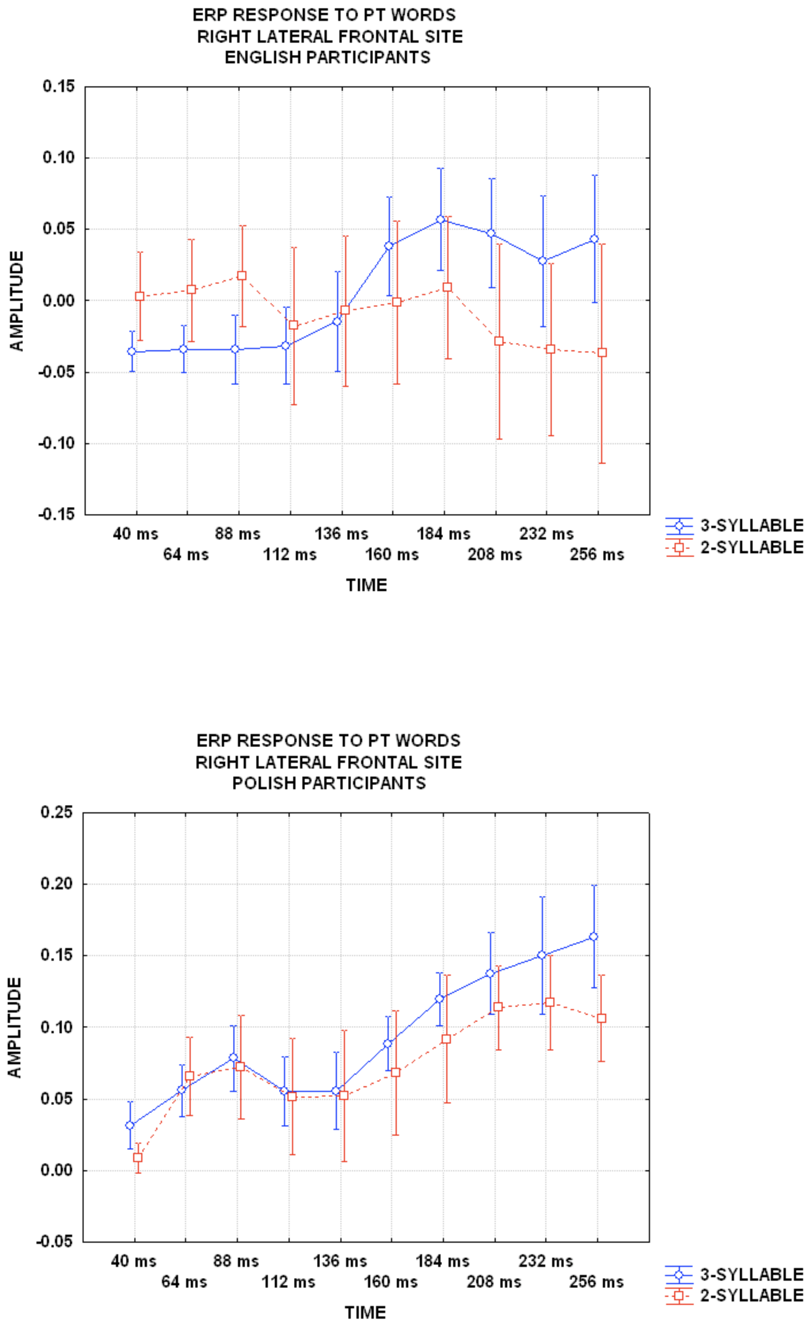


Figure 3.9 depicts the English (top) and Polish listeners' (bottom) response to the *pt* words from the right lateral frontal site. Notice the small variability for the Polish participants' response to the 2-syllable words between 208 and 256 ms relative to the English participants' response

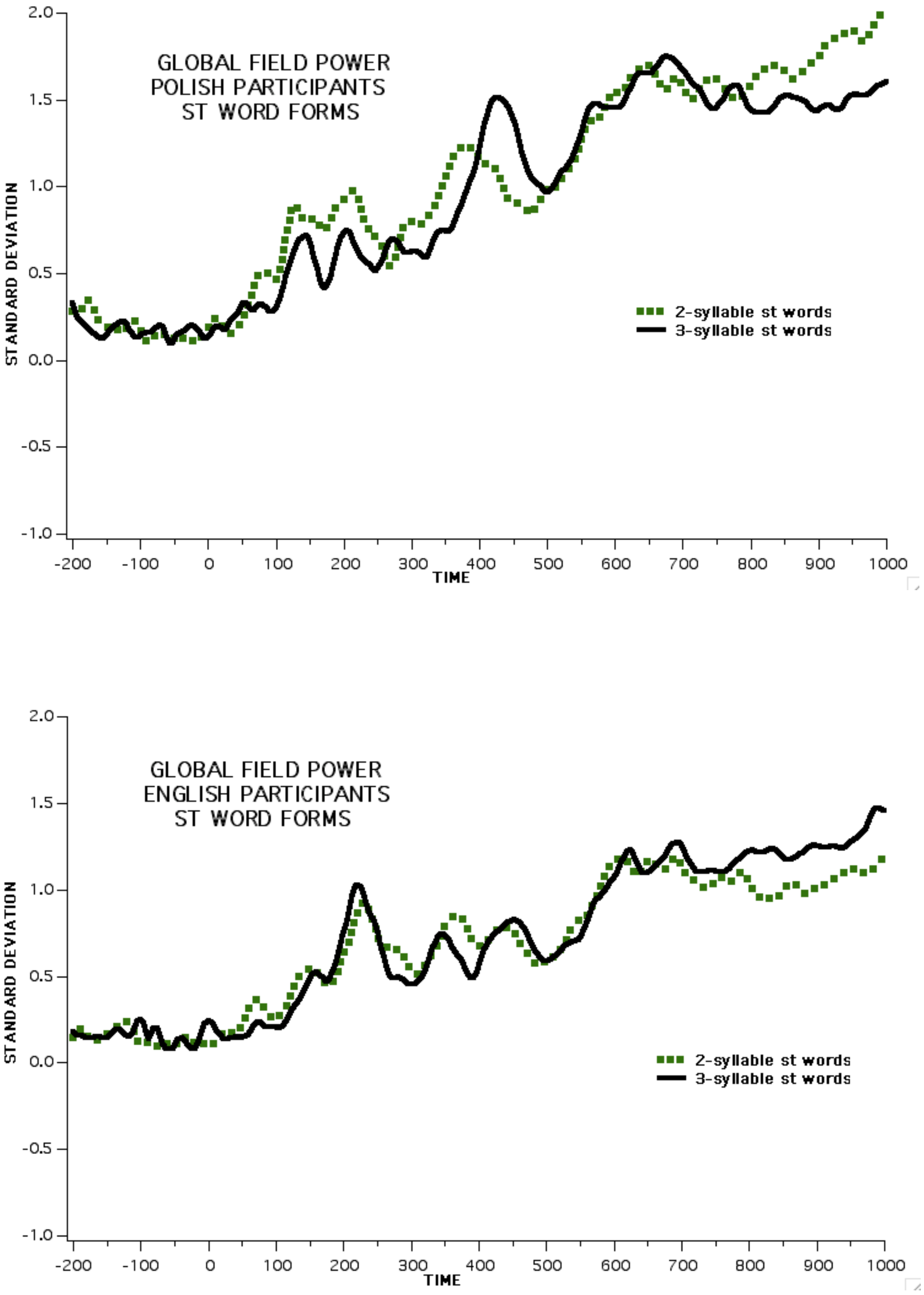
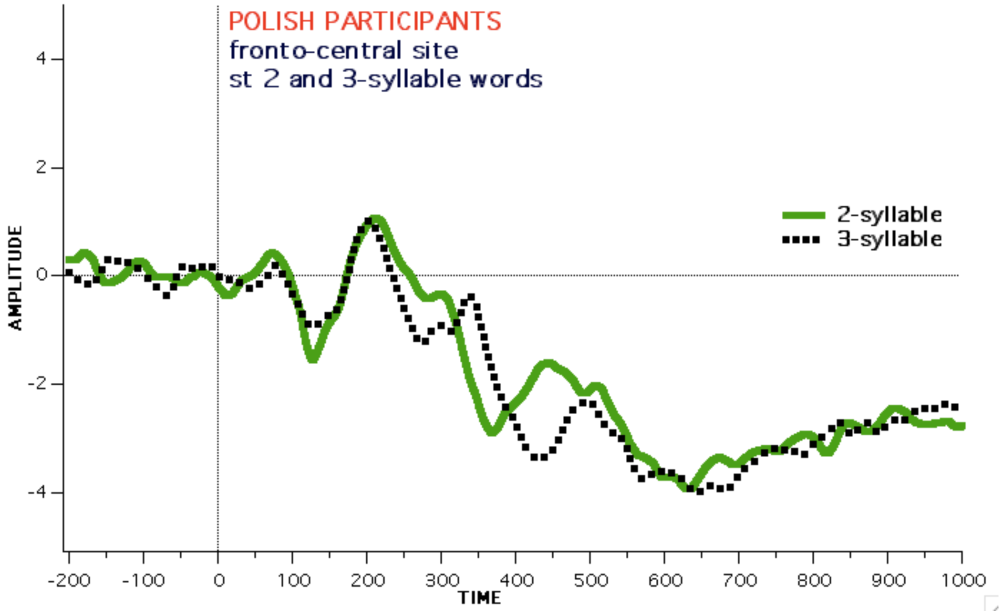


Figure 3.10 illustrates the GFP analysis of the response to 2 and 3 syllable *st* words for Polish (top) and English (bottom) listeners



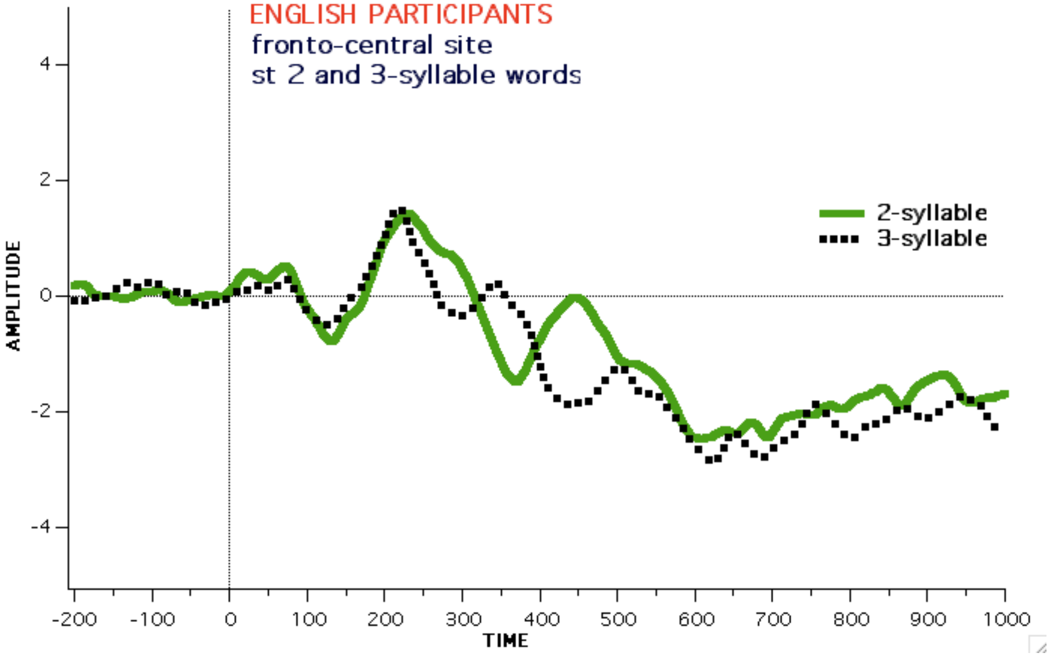


Figure 3.11 depicts the grand mean auditory word signature waveforms in response to the 2 and 3-syllable *st* words from the fronto-central site for the Polish (top) and English (bottom) participant group. Notice the highly similar auditory word signature patterns for the 2 and 3-syllable words for the language groups. The raw voltage waveforms are shown

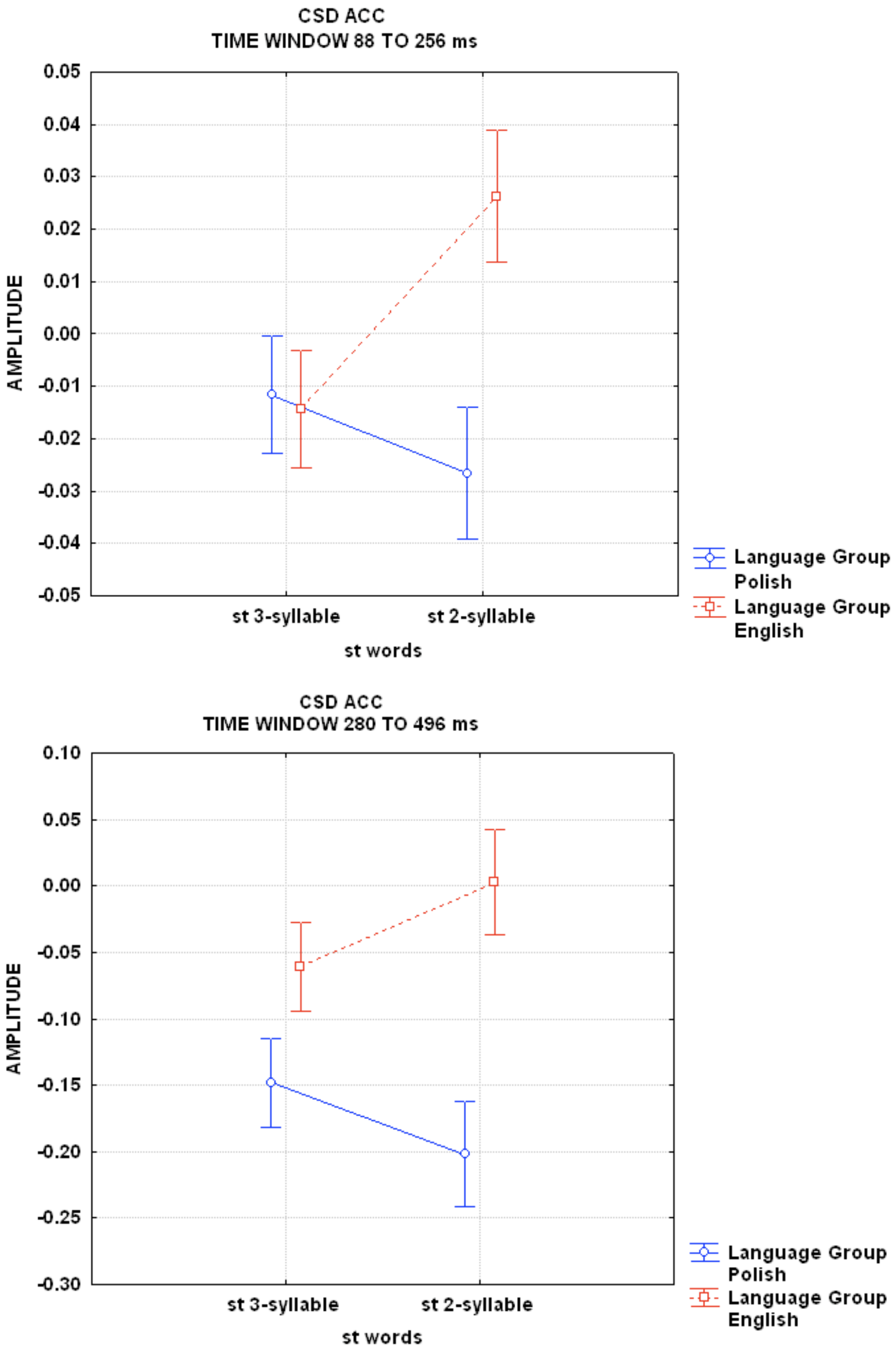


Figure 3.12 depicts the different language group P1N1P2 response (CSD) to the 2-syllable *st* words for the time interval between 88 and 256 ms (top) and for the time interval between 280 and 496 ms (bottom). Note the different scales

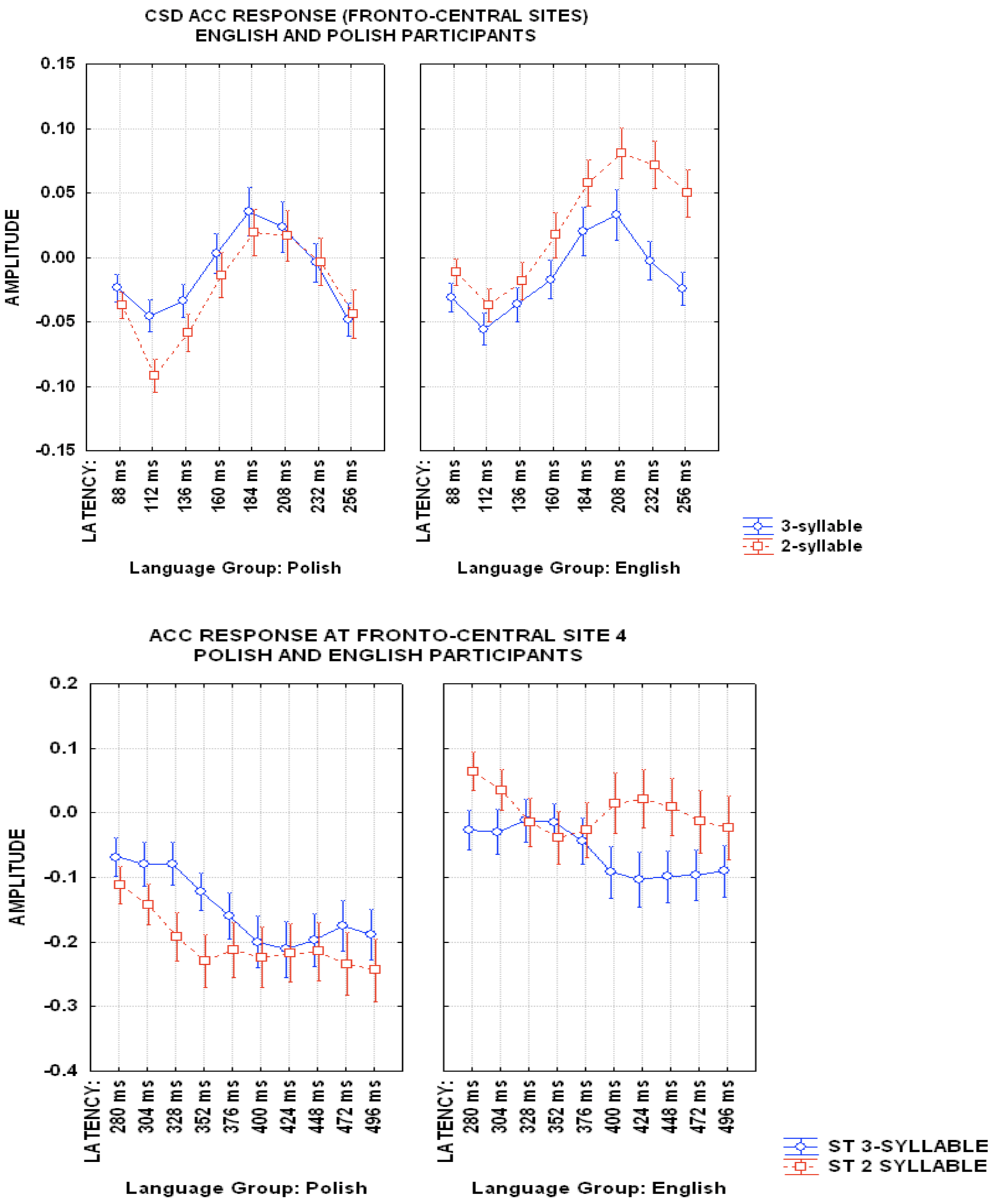


Figure 3.13 (top) depicts the P1N1P2 response (CSD) to the 2 and 3-syllable *st* words between 88 and 256 ms from four fronto-central sites. Notice the negative response to the 2-syllable words relative to the 3-syllable words for the Polish group and a reverse pattern for the English group. (bottom) A similar pattern is seen for the later intervals between 280 and 496 ms from the midline fronto-central site

CHAPTER 4

CONCLUSION

The phonotactic patterns of Polish and English permit a natural experiment, which allowed us to investigate the effect of exposure (or lack of) to a word onset cluster (/pt/) on speech perception and on various stages of neural speech processing. The results of the dissertation have demonstrated that (1) contextual features of the phonemes are intrinsic to perceptual speech processing and (2) native-language patterns of speech perception are reflected in the ERP waveform in early and late stages of neural speech processing. The results revealed that the acoustic signals of speech are registered during early stages of processing within cortical sources underlying the P1N1P2 complex whether or not the acoustic features of phonemes are perceived and native-language speech perception is reflected within the T-Complex from lateral-temporal sites at early stages of processing beginning at 40 ms.

The current dissertation investigated native-Polish and English listeners' behavioral and ERP responses to word onset consonant clusters (/st/ and /pt/) and their contrast segments (/set/ and /pet/). The phonotactic pattern of English, which precludes the word onset /pt/ cluster, resulted in English listeners being unable to distinguish 2 and 3-syllable *pt* words for a syllable identification task. These same English and Polish participants responded to the acoustic signals of the 2 and 3-syllable *pt* words at early levels of neural processing within the P1N1P2 complex in the same manner. Therefore, the underlying sources of the P1N1P2 sensory-obligatory component register the acoustic signals of speech whether or not they are perceived. This finding is consistent with studies of speech perception, which have found that individuals retain the ability to discriminate non-native contrasts under controlled laboratory conditions (Strange & Shafer, 2008, Strange & Dittman, 1984). Furthermore, our results revealed that both the acoustic and linguistic aspects of speech are registered at early stages of processing but from different brain sources. The current results provide neurophysiological support for Strange & Shafer's automatic selective perception model (2008), which states that native (phonological contrasts) and non-native (acoustic contrasts not phonological in the native-language) contrasts can be discriminated but in a different manner. Strange & Shafer reported that native-language contrasts are perceived rapidly and automatically even under suboptimal conditions and that non-native contrasts can be perceived but in a different (non-automatic) manner. Japanese listeners trained to accurately discriminate /r/ and /l/ within word onset (i.e., rock/lock) using a synthetically created continuum were unable to generalize learning to other words (i.e., rake/lake) or contexts (Strange & Dittman, 1984). In summary, the dissertation findings suggest a parallel cortical network for perceptual processing of linguistic discriminations (phonologically contrastive within the native-language) that begins at early stages of cortical speech processing consistent with the view that a different processzz occurs for perception of native (linguistic) and non-native contrasts (acoustic).

Martin, Tremblay & Korczak (2008) suggested that the P1N1P2 complex provides information concerning an individual's access to the acoustic signals of speech. Our findings together with those of Strange & Shafer (2008) support Martin et al., in that the acoustic signals of speech are registered at cortical levels irrespective of native-language input. The non-native listeners have access to the acoustic information as demonstrated under controlled laboratory conditions. The neural networks for automatic speech perception follow a different path, thereby affecting behavioral speech perception within more natural speech environments. Davis and Zerlin (1966) proposed that the sources underlying the fronto-central P1N1P2 complex do not precede cortical processes underlying the perception of speech in a serial manner (Naatanen & Winkler, 1999) consistent with the current findings.

ERP epochs time-locked to both prime and target word stimuli revealed native-language speech perception to be reflected during early and late stages of neural processing. Epochs time-locked to the target words (reported in Chapter 2) revealed speech perception to be reflected at late stages of neural processing. The behavioral task required that the participants determine the number of syllables within the target words and the LPC and the P400 response reflected the behavioral response of the native-language groups. ERP epochs time-locked to the prime words, which showed clear sensory obligatory components (ITI 2 ms), revealed the native-language influence to be reflected at early stages of neural speech processing (reported in Chapter 3). An overt response was not required of the prime word stimuli and attention was likely drawn to the auditory modality during the experiment by the nature of the behavioral task. Factors of attention may have resulted in language group differences within the P1N1P2 complex to a phonological contrast (i.e., *st* contrast) present in both Polish and English. To our knowledge, an investigation of sensory obligatory component responses to specific sound contrasts within two native-language groups has not been previously demonstrated. The current results have implications for future studies with individuals having speech and language impairment. Comparison of children having typical or atypical phonological and language development using the current experimental design will determine whether the linguistic impairment has underlying involvement of early or late stages of cortical processing and whether the linguistic impairment has underlying involvement of networks processing acoustic or linguistic aspects of the speech stimuli.

In chapter 2, we demonstrated that the acoustic features of phonemes that change with context are intrinsic to native-language speech perception. We argued that the acoustic features of a phoneme sequence that form a word unit consist of a unique pattern, which can explain findings concerning word recognition in the both developmental and adult literature and is consistent with other neurophysiological studies of speech processing Steinschneider, Fishman & Arezzo, 2003; Steinschneider, Schroeder, Arezzo & Vaughan, 1995). Saffran, Aslin and Newport (1996) argue that infants have an inborn ability to detect patterns from the linguistic environment. The current research has demonstrated that the contextual (acoustic) features of phonemes are intrinsic to perceptual speech processing and, therefore, result in a unique pattern for each phoneme sequence. We found perception of linguistic contrasts, which include these contextual (acoustic) features to be reflected within early stages of cortical speech processing in adult subjects. Infants may encode the acoustic features of phonemes within the native-language input within cortical auditory pathways, which would permit rapid detection of word-like units and awareness of phonotactic patterns of the native-language suggesting that detection of patterns (transitional probability learning) is not the processes by which word units are detected at these early stages of development. In future research, infant developmental changes pertaining to neural correlates of early and late stage acoustic feature processing will be investigated. A first step would be to determine whether the acoustic signals of the 2 and 3-syllable *pt* nonsense words distinguished within the adult P1N1P2 complex using the current experimental design are evident in infants' sensory obligatory ERP components. The current results found that the *pt* contrast, which is a native-language contrast for only the Polish group, did not modulate the sensory-obligatory P1N1P2 complex in adults. On the other hand, we found that native-language sound sequences present in Polish and English (*st* stimuli) affected the fronto-central response (P1N1P2 complex) differently in the language groups. Non-native contrasts perceived during early infancy are perceived more poorly by 12 months of age (Strange & Shafer, 2008; Kuhl, 2000). Hence, we question whether exposure (or lack of) to phonotactic sound sequences within the native-language affect the P1N1P2 complex differently at various developmental stages. Also, we argued that differing levels of attention to the *st* stimuli by the English and Polish listeners might have affected the P1N1P2 complex in response to the *st* 2 and 3-syllable words, leading us to question the role of attention in the infants' development of speech and language.

The contrast of the /s/ and /p/ within the word onset sequences resulted in significant effects within the ERP components including presence or absence of components (e.g., Tb component target word data), latency (TN component target word) and topographical differences (P400 CSD maps target word) and hemisphere effects to the contrast (target word TN, prime word date). These results suggest that studies of word processing should consider the specific sounds and features within words because they affect neural speech processing. Whether similar results will occur using real word stimuli rather than nonsense word stimuli awaits future research. Different results in response to the *st* and *pt* stimuli also have implications for individuals having speech and language impairment. For example, future examination of children having typical or atypical phonological and language development using the current experimental design can address issues regarding the manner of impairment, whether spectral or temporal.

In conclusion, the dissertation revealed that contextual features of phonemes are intrinsic to perceptual speech processing and native-language experiences modulate early and late stages of speech processing. At early stages, native-language experiences modulated the T-Complex response but not the P1N1P2 complex response suggesting different neural networks for perceptual speech processing and acoustic processing.

REFERENCES

Alho, K., Connolly, J., Cheour, M., Lehtokoski, A., Huotilainen, M., Virtanen, J., Aulanko, R. & IImoniemi, R. (1998). Hemisphereic lateralization in preattentive processing of speech sounds. *Neuroscience Letters*, 258, 9-12.

Andruski, J., Blumstein, S. & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition*, 52, 163-187.

Belin, P., Zatorre, RJ., Lafaille, P., Ahad, P. & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403, 302-312.

Berent, I., Steriade, D., Lennertz, T. & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition,* 104, 591-630.

Bertoncini, J., Bijeljac-Babic, R., Kennedy, L., Jusczyk, P. & Mehler, J. (1988). An investigation of young infants' perceptual representations of speech sounds. *Journal of Experimental Psychology: General*, 117 (1), 21-33.

Bertoncini, J. & Mehler, J. (1981). Syllables as units in infant speech perception. *Infant Behavior and Development*, 4, 247-260.

Birdsong, D. (2005). Why not fossilization. In Z.H. Hahn & T. Odin (Eds.), *Studies of Fossilization: Second Language Acquisition* (Ch. 9 pp. 173-188) Clevedon: Multilingual Matters.

Bordon, G., Harris, K. & Raphael, L. (2003). *Speech science primer: physiology, acoustics, and perception of speech (4th Ed.)*. Baltimore, MD: Lippincott, Williams & Wilkins.

Burnham, D.K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics*, 7, 207-240.

Ceponiene, R., Alku, M., Westerfield, M., Torki, M. & Townsend, J. (2005). ERPs differentiate syllable and nonphonetic sound processing in children and adults. *Psychophysiology*, 42, 391-406.

Connolly, J. & Phillips, N. (1994). Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *Journal of Cognitive Neuroscience*, 6 (3), 256-266.

Crowley, K. & Colrain, I. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology*, 115 (4), 732-744.

Davidson, L., Shaw, J. & Adams, T. (2007). The effect of word learning on the perception of non-native consonant sequences. *Journal of Acoustical Society of America*, 122 (6), 3697–3709.

Davidson, L. & Stone, M. (2004). Epenthesis versus gestural mistiming in consonant cluster production: An ultrasound study. In M. Tsujimura & G. Garding (Eds.), (WCCFL) 22 *Proceedings of the West Coast Conference on Formal Linguistics,* 165-178.

Dehaene-Lambertz, G., Dupoux, E. & Gout, A. (2000). Electrophysiological correlates of phonological processing: A cross-linguistic study. *Journal of Cognitive Neuroscience*, 12 (4), 635-647.

Dupoux, E., Hirose, Y., Kakahi, K., Pallier, C. & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25 (6), 1568-1578.

Eggermont, J (2007). Electrical and magnetic fields of synchronous neural activity. In R. Burkard, M. Don & J. Eggermont (Eds.). *Auditory Evoked Potentials: Basic Principles and Clinical Application,* New York: Lippincott, Williams & Wilkins.

Eggermont, J. (2005). Correlated Neural Activity: Epiphenomenon or part of the neural code? In R. Konig, P. Heil, E. Budinger & H. Scheich (Eds.). *The Auditory Cortex: A synthesis of human and animal research,* Mahwah, NJ: Lawrence Erlbaum Associates, Publ.

Ferree, T.C., Luu, P., Russell, G.S. & Tucker, D.M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology,* 112, 536-544.

Fisiak, J., Adamska-Sa?aciak, A., Ga;siorowski, P. & Jankowski, M. (eds.). (2003). *The new English-Polish and Polish-English Kosciuszko Foundation dictionary*. Kraków: Universitas.

Friedrich, C.K., Schild, U. & Roder, B., (2009). Electrophysiological indices of word fragment priming allow characterizing neural stages of speech recognition. *Biological Psychology*, 80 (1), 105-113.

Frishkoff, GA., Tucker, DM., Davey, C. & Scherg, M. (2004). Frontal and posterior sources of event-related potentials in semantic comprehension. *Cognitive Brain Research*, 20, 329-354.

Gierut, J. (1999). Syllable onsets: Clusters and Adjuncts in Acquisition. *Journal of Speech, Language, and Hearing Research*, 42, 708-726.

Halicki, S. (2010). *Learner Knowledge of Target Phonotactics: Judgements of French Word Transformations* (In Lincom Studies in Language Acquisition 27). Europa: Lincom.

Halle, P., Segui, J., Frauenfelder, U. & Meunier, C. (1998). Processing of illegal consonant clusters: A case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance*, 24 (2), 592-608.

Hestvik, A., Maxfield, N., Schwartz, R. & Shafer, VL. (2007). Brain responses to filled gaps. *Brain and Language*, 100 (3) 301-316.

Holcomb, P, Grainger, J. & O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*, 14 (6), p 938-950.

Huttenlocher, P. & Dabholkar, A. (1997). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of Comparative Neurology*, 387, 167-178.

Isreal, J.B., Chesney, G.L., Wickens, C.D. & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology*, 17 (3), 259-273.

Jusczyk, P.W. (2000). *The discovery of spoken language*. Cambridge, MA: MIT Press

Jusczyk, P.W., Friederici, A.D., Wessels, J.M., Svenkerud, V.Y. & Jusczyk, A.M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language*, 32, 402-420.

Jusczyk, P., Luce, P. & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language*, 33, 630-645.

Kaukoranta, E., Hari, R. & Lounasmaa, O. (1987). Responses of the human auditory cortex to vowel onset after fricative consonants. *Experimental Brain Research*, 69, 19-23.

Kuhl, P. (2000). A new view of language acquisition. *PNAS,* 97(22), 11850-11857.

Ladefoged, P. (2001). *A Course in Phonetics* (Fourth Edition). United States: Thomson Learning, Inc.

Lehmann, D., & Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and clinical neurophysiology,* 48, 609-621.

Leuthold, H. & Sommer, W. (1998). Postperceptual effects and P300. *Psychophysiology,* 35, 34-46.

Luce, P., Goldinger, S., Auer, E. & Vitavitch, M. (2000). Phonetic priming, neighborhood activation and PARSYN. *Perception and Psychophysics,* 62, 615-625.

Luck, S. (1998). Sources of dual-task interference: Evidence from human electrophysiology. *Psychological Science*, 9 (3), 223-227.

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique.* Cambridge, MA: The MIT Press.

Martin, B., Tremblay, K. & Korczak, P. (2008). Speech evoked potentials: From the laboratory to the clinic. *Ear and Hearing*, 29 (3), 285-313.

Martin, B., Shafer, V. L., Morr, M., Kreuzer, J. & Kurtzberg, D. (2003). Maturation of Mismatch Negativity: A scalp current density analysis. *Ear and Hearing*, 24 (6), 463-471.

Martin, B., Tremblay, K. & Stapells, D. (2007). Principles and Application of Cortical Auditory Evoked Potentials. In R. Burkard, M. Don, & J. Eggermont (Eds.) *Auditory Evoked Potentials: Basic Principles and Clinical Application*, New York: Lippincott, Williams & Wilkins.

McCallum, W. C., Farmer, S. F. & Pocock, P.V. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and clinical neurophysiology*, 59, 477 - 488.

McCarthy, G. & Donchin, E. (1981). A metric for thought: A comparison of P300 as a measure of stimulus evaluation time. *Science*, 211, 77-80.

Moats, L. (2000). Speech to Print. Baltimore, MD: Brookes.

Naatanen, R. & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, 24 (4), 375-425.

Nobre, A. & McCarthy, G. (1994). Language-related ERP's: Scalp distributions and modulation by word type and semantic priming. *Journal of Cognitive Neuroscience*, 6(3), 233-255.

O'Rourke, T. & Holcomb, P. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological Psychology*, 60, 121-150.

Ostroff, J., Martin, B. & Boothroyd, A. (1998). Cortical evoked response to acoustic change within a syllable. *Ear and Hearing,* 19 (4), 290-297.

Pernier, J., Perrin, F. & Bertrand, O. (1988). Scalp current density fields; Concept and properties. *Electroencephalography and clinical neurophysiology*, 69, 385 - 389.

Picton, T. (1992). The P300 wave of the human event-related potential. *Journal of clinical neurophysiology; Official publication of the American Electroencephalographic Society*, 9, 456-479.

Picton, T. & Hillyard, S. (1074). Human auditory evoked potentials. II Effects of attention. *Electroencephalograpy and clinical neurophysiology*, 36, 191-199.

Picton, T., Linden, R., Hamel, G. & Maru, J. (1983). Aspects of Averaging. *Seminars in Hearing*, 4 (4), 327-340.

Ponton, C., Eggermont, J., Khosla, D., Kwong, B. & Don, M. (2002). Maturation of human central auditory system activity: separating auditory evoked potentials by dipole source modeling. Clinical Neurophysiology, 113, 407-420.

Picton, T., Linden, R., Hamel, G. & Maru, J. (1983). Aspects of Averaging. *Seminars in Hearing*, 4 (4), 327-340.

Praamstra, P. & Stegeman, D. (1993). Phonological effects on the auditory N400 event-related brain potential. *Cognitive Brain Research*, 1(2), 73-86.

Praamstra, P., Meyer, A. & Levelt, W. (1994). Neurophysiological manifestations of phonological processing: Latency variation of a negative ERP component time-locked to phonological mismatch. *Journal of Cognitive Neuroscience*, 6(3), 204-219.

Sebastian-Galles, N., Rodriguez-Fornells, A., de Diego-Balaguer, R. & Diaz, B. (2006). First-and Second-language phonological representations in the mental lexicon. Journal of Cognitive Neuroscience, 18 (8) p 1277-1291.

Picton, T.W. (1992). The P300 wave of the human event-related potential. *Journal of clinical neurophysiology; Official publication of the American Electroencephalographic Society*, 9, 456-479.

Picton, T.W. & Hillyard, S.A. (1974). Human Auditory Evoked Potentials II: Effects of Attention. *Electroencephalography & Clinical Neurophysiology*, 36, 191-199.

Rinne, T., Alho, K., Alku, P., Holi, M., Sinkkonen, J., Virtanen, J., Bertrand, O. & Naatanen, R. (1999). Analysis of speech sounds is left-hemisphere predominant at 100-150 ms after sound onset. *NeuroReport Letters*, 10, 1113-1117.

Saffran, J. R. Aslin, R. & Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.

Saffran, J.R., Newport, E.L. & Aslin, R. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606-621.

Schwartz, R., Maxfield, N., Gross, M. & Shafer, V.L. (in preparation).

Shafer, V.L., Ponton, C., Datta, H., Morr, M. & Schwartz, R. (2007). Neurophysiological indices of attention to speech in children with specific language impairment. *Clinical Neurophysiology*, 118, 1230-1243.

Shafer,V.L., Schwartz, R. & Kessler, K. (2002). ERP Indices of phonological and lexical processing in children and adults. In BUCLD 27, Proceedings, 762-774.

Shafer, V.L., Schwartz, R. & Kurtzberg, D. (2004). Language-specific memory traces of consonants in the brain. *Cognitive Brain Research*, 18, 242-254.

Shafer, V.L., Schwartz, R., & Martin, B. (in press). Evidence of Deficient Central Speech Processing in Children with Specific Language Impairment: the T-Complex.

Sheehan, K., McArthur, G., Bishop, D. (2005). Is discrimination training necessary to cause changes in the P2 auditory event-related brain potential to speech sounds. *Cognitive Brain Research*, 25, 547-553.

Siegel, S. & Castellan, N. (1988). *Nonparametric Statistics for The Behavioral Sciences.* New York, McGraw-Hill Publisher.

Sommer, W. & Matt, J. (1990). Awareness of P300-related cognitive processes: A signal detection approach. *Psychophysiology*, 27(5), 575-585.

Squires, KC., Hillyard, S. A. & Lindsay, P.H. (1973). Vertex potentials evoked during auditory signal detection; relation to decision criteria. *Perception & Psychophysics,* 14, 265-272.

Steinschneider, M. & Dunn, M. (2002). Electrophysiology in developmental neurophyschology. In S.J. Segalowitz & I. Rapin (Eds.), *Handbook of Neuropsychology 2nd Edition* (Vol.8, Part 1, Chap. 5). New York: Elsevier Science B.V.

Steinschneider, M., Fishman, Y. & Arezzo, J. (2003). Representation of voice onset time (VOT) speech parameter in population responses within primary auditory cortex of the awake monkey. *Journal of the Acoustical Society of America,* 114 (1), 307-321.

Steinschneider, M., Kurtzberg, D., & Vaughan Jr., H. (1992). Event-related potentials in developmental neuropsychology. In I. Rapin & S.J. Segalowitz (Eds.), *Handbook of Neurophysiology* (Vol.6, 239-300). New York: Elsevier.

Steinschneider, M., Schroeder, C., Arezzo, J., & Vaughan, Jr., H. (1995). Physiologic correlates of the voice onset time boundary in primary auditory cortex (A1) of the awake monkey: Temporal response patterns. *Brain and Language*, 48, 326-340.

Steinschneider, M., Volkov, I., Fishman, Y., Oya, H., Arezzo, J., & Howard III, M. (2005). Intracortical responses in human and monkey primary auditory cortex support a temporal processing mechanism for encoding of the voice onset time phonetic parameter. *Cerebral Cortex*, 15, 170-186.

Strange, W., & Shafer, V.L. (2008). Speech perception in second language learners: The re-education of selective perception. In J.G. Hansen Edwards & M.L. Zampini (Eds.) *Phonology and Second Language Acquisition.* Philadelphia, PA: John Benjamins Publ Co.

Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J. & Lawson, K. (2008). The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hearing Research,* 236, 61-79.

Tonnquist-Uhlen, I., Ponton, C., Eggermont, J., Kwong, B. & Don, M. (2003). Maturation of human central auditory system activity: the T-complex. *Clinical Neurophysiology,* 114, 685-701.

Tremblay, K. & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech, Language, and Hearing Research*, 45, 564-572.

Vitavitch, M. (2007). The spread of the phonological neighborhood influences spoken word recognition. *Memory & Cognition*, 35, 166-175.

Vitevitch, M.S. & Luce, P.A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. Behavior Research Methods, Instruments, and Computers, 36, 481-487. [PDF version]

Weber, A. & Cutler, A. (2006). First-language phonotactics in second-language listening. *Journal of Acoustical Society of America*, 119 (1), 597-607.

Winkler, I., Czigler, I., Jaramillo, M., Paavilainen, P. & Naatanen, R. (1998). Temporal constraints of auditory event synthesis: Evidence from ERPs. *NeuroReport*, 9, 495-499.

Wolpaw, JR. & Penry, JK. (1975). A temporal component of the auditory evoked response. *Electroencephalography and clinical neurophysiology,* 39, 609-620.

Wunderlich, J. & Cone-Wesson, B. (2001). Effects of stimulus frequency and complexity on the mismatched negativity and other components of the cortical ausitory-evoked potentials. *Journal of Acoustical Society of America*, 109 (4), 1526-1537.

Zhang, Y., Kuhl, P., Imada, T., Kotani M. & Tohkura, Y. (2005). Effects of language experience: Neural commitment to language-specific auditory patterns. *NeuroImage*, 26, 703-720.