Brain-Based Individual Difference Measures of Reading Skill in Deaf and Hearing Adults

Alison S. Mehravari

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:

Lee Osterhout, Chair

Chantel Prat

David Perkel

Program Authorized to Offer Degree:

Neuroscience

University of Washington

Abstract

Brain-Based Individual Difference Measures of Reading Skill in Deaf and Hearing Adults

Alison S. Mehravari

Chair of the Supervisory Committee:
Professor Lee Osterhout
Department of Psychology, Program in Neuroscience

A majority of deaf students leave high school reading at or below a fourth grade level, but some deaf individuals do become highly proficient readers. There is disagreement about the causes of this reading difficulty, and by association, disagreement about the effectiveness of different strategies for teaching reading to deaf children. The goal of this study was to use real-time measures of neural language processing to better assess if deaf and hearing adults read proficiently in similar or different ways. Hearing native English speakers and non-native signing deaf adults read English sentences and word pairs while event-related potentials (ERPs) were recorded. The magnitude of ERP responses was compared to participants' standardized reading comprehension test scores. The best deaf readers had the largest responses to information about meaning in sentences, while the best hearing readers had the largest responses to information about grammar in sentences. These results show that equally proficient hearing and non-native signing deaf adults read in different ways, and suggest that for deaf individuals, the most important aspect of successful reading instruction may be increasing vocabulary knowledge. These results also provide strong support for the feasibility of individual differences analysis of ERP data, especially in highly variable populations.

Table of Contents

List of Figures

List of Tables

Chaper 1. Introduction

Reading can be difficult for many people who are deaf. Numerous studies have shown that the median reading level of deaf students graduating from high school in the United States is a fourth grade level (Allen, 1986; Qi & Mitchell, 2012; Trybus & Karchmer, 1977). Though this difficulty has been recognized for half a century, the reading achievement of deaf students has changed little during that time (Qi & Mitchell, 2012). Nevertheless, some deaf individuals do become proficient readers. Approximately 10% of deaf students read above an eighth grade level (Qi & Mitchell, 2012; Traxler, 2000). This shows that, though many deaf students struggle to read, it is possible for someone who is deaf to become a proficient reader (Goldin-Meadow & Mayberry, 2001). To improve the potential for all deaf individuals to read well, we must determine what allows some to become proficient readers, while many others struggle (Mayberry, del Giudice, & Lieberman, 2011).

1.1. Hypotheses about deaf literacy challenges. Historically, there have been two main hypotheses about the causes of reading difficulty in deaf individuals.

1.1.1. Phonology hypothesis. The first hypothesis is that reading difficulties arise because deaf readers cannot directly access sound information of the language being read. For hearing children, learning to read involves learning the associations between the spoken language they know and the written form of that language, a task for which phonological knowledge is central (Goldin-Meadow & Mayberry, 2001; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2001). This first hypothesis posits that phonological knowledge is as important for deaf children learning how to read as it is for hearing children, and that reading difficulties in deaf individuals stem from challenges in forming phonological representations (Hanson, 1989; Mayer & Trezek, 2014; Paul, Wang, Trezek, & Luckner, 2009; Wang, Trezek, Luckner, & Paul, 2008). Though deaf individuals do not hear the language they are reading, they can learn about phonology of a spoken language via alternate routes, including speechreading, articulatory feedback during speech, and systems such as Cued Speech and Visual Phonics, which use hand cues around the mouth to represent consonants and vowels during speech (Wang et al., 2008). Numerous studies have shown that some deaf individuals use phonological information while reading (for reviews see Perfetti & Sandak, 2000; Wang et al., 2008). For example, deaf individuals have been shown to judge at above-chance levels whether words rhyme (in a task in which orthography is controlled) (Hanson & Fowler, 1987), perform picture rhyming tasks at an above-chance level (Colin, Magnan, Ecalle, & Leybaert, 2007), and show similar interference effects of phonology as hearing readers do in reading tasks involving tongue-twisters or other types of phonological interference (Perfetti & Sandak, 2000). Many of these studies find that more phonological knowledge is correlated with better reading ability (Colin et al., 2007; LaSasso, Crain, & Leybaert, 2003; Luetke-Stahlman & Nielsen, 2003). Proponents of the importance of phonology hypothesis emphasize that teaching phonological awareness skills is crucial in order for deaf children to learn to read successfully.

1.1.2. Language proficiency hypothesis. The second hypothesis is that reading difficulties arise because many deaf children are not proficient in any language, signed or spoken, when they begin to learn to read. Hearing children, when exposed to a spoken language such as English early in life, learn spoken language relatively automatically. Deaf children, by the nature of not hearing, do not learn spoken English as naturally as hearing children do. However, when raised in a sign language-rich environment from birth, deaf children learn a sign language as easily as hearing children learn a spoken language (Meadow-Orlans, Spencer, & Koester, 2004; Meier & Newport, 1990; Sandler & Lillo-Martin, 2001) Deaf children acquiring a sign language from birth pass through similar language-acquisition milestones and at similar rates as do hearing children acquiring a spoken language (Sandler & Lillo-Martin, 2001).

A few important notes about sign languages, and American Sign Language (ASL) in particular, are warranted. First, natural sign languages are not based off of the spoken languages of the region (Bellugi & Studdert-Kennedy, 1980; Goldin-Meadow & Mayberry, 2001; Klima & Bellugi, 1979; Lane & Grosjean, 1980). ASL is not simply a signed form of English; it is a completely different language from English in the same way that Spanish is a different language than English; it just happens that ASL is a signed language rather than a spoken language. The grammar of ASL is different than the grammar of English, and the autonomy of signed languages is made clear by the fact that American Sign Language is quite distinct from British Sign Language (Goldin-Meadow & Mayberry, 2001).

In discussing sign languages as languages, it is important to define what a natural language is. Though scholars sometimes differ in the exact definition of a natural language, the main principle is that natural languages develop spontaneously wherever there are people who have opportunities to interact and communicate, without being consciously invented as a system of communication (Lyons, 1991; Sandler & Lillo-Martin, 2001). Natural sign languages like ASL allow for the same depth of communication and have equivalent linguistic complexity as natural spoken languages do. As Sandler and Lillo-Martin so clearly explain, “Sign language can “do” everything that spoken language can” (Sandler & Lillo-Martin, 2001). Though historically there has been some debate about whether sign languages are natural languages, scholars have reached general agreement that that naturally occurring sign languages such as ASL are indeed natural languages (Friedmann & Szterman, 2011; Lederberg, Schick, & Spencer, 2013; Sandler & Lillo-Martin, 2001; Stokoe, Casterline, & Croneberg, 1965).

Natural languages are distinguished from artificially created languages, such as programming languages. Specifically, natural sign languages are distinguished from created manual communications systems such as Signed Exact English (SEE) and other forms of Manually Coded English (MCE). Manually Coded English systems use manual (signed) communication to exactly replicate the grammar and vocabulary of spoken English (Allen et al., 2009; Emmorey, 2001; Goldin-Meadow & Mayberry, 2001; Sandler & Lillo-Martin, 2001). While incorporating some signs from ASL, Manually Coded English systems also include handshapes and signs for words that do not exist in ASL (i.e., different conjugations of a verb (e.g. *was* and *were*) which would not be differentiated in ASL) as well as signs for English word affixes like *–ing and –ed,* which again are not a part of ASL (Emmorey, 2001; Klima & Bellugi, 1979). As Emmorey describes, "SEE was created by a committee" – it is not a natural language (Emmorey, 2001). In addition to SEE and other forms of Manually Coded English, another form of manual communication that is not a true sign language is Pidgin Sign[ed] English (PSE), also known as contact sign, which uses ASL signs in an English word order, but without so many of the English word affixes used in SEE (Emmorey, 2001; Reilly & McIntire, 1980; Woodward, 1973). PSE is often used (intentionally or unintentionally) by people whose first language is English – i.e., hearing non-native signers (Emmorey, 2001).

Understanding the differences between ASL and other forms of manual communication is important for understanding the wide variety of language backgrounds deaf children can grow up in. As has been mentioned, when deaf children are raised in a natural sign language-rich environment from birth, they will learn that sign language as easily as hearing children learn a spoken language (Meadow-Orlans et al., 2004; Meier & Newport, 1990). Research has shown that in contrast to acquisition of a natural sign language, forms of Manually Coded English are not acquired as easily by deaf children. Studies in deaf children who had only been exposed to Signed Exact English from a young age showed that those children experienced difficulty in acquiring SEE, and when using it, often modified it in ways that made it more similar to natural sign languages, i.e., by adding more spatial aspects to the signs (Supalla, 1991). In ASL, spatial aspects of signs or hand movements are used to convey grammatical information; in Manually Coded English systems, they do not (Supalla, 1991).

Approximately ninety percent of deaf children have hearing parents (Goldin-Meadow & Mayberry, 2001), and many are raised with a mixture of spoken and manual communication that is not a natural signed language: SEE, PSE, and/or home sign systems, any of which may be used alone or in conjunction with spoken English or the native spoken language of the parents (Easterbrooks & Beal-Alvarez, 2013). Though these children may become familiar with spoken English through speechreading or forms of Manually Coded English, the process of learning a spoken language is not automatic for deaf children in the same way that it seems to be for hearing children. Because of this, many deaf children do not become proficient in any natural language before they begin to learn to read (Emmorey, 2001; Goldin-Meadow & Mayberry, 2001). However, a small proportion of deaf children, most born to deaf parents, are raised in a sign language-rich environment from an early age. When deaf children raised learning a signed language begin to learn to read English, though they do not know English, they are proficient in a natural language, have a basic understanding of the concepts of grammar and vocabulary, and can use fingerspelling to aid word learning (Lederberg et al., 2013). The second hypothesis posits that proficiency in any language is most important for learning to read, because this allows children to make connections between the language they know and the language they are learning to read (Goldin-Meadow & Mayberry, 2001). In support of this hypothesis, a number of studies find that the best predictor of reading skills in deaf children is skill in a signed language (Chamberlain & Mayberry, 2008; Hermans, Knoors, Ormel, & Verhoeven, 2008; Padden & Ramsay, 2000; Strong & Prinz, 2000). Notably, children who improve their knowledge of a form of Manually Coded English but not of ASL do not show an associated improvement in reading skill (Goldin-Meadow & Mayberry, 2001).

1.1.3. Which hypothesis is it? Or is it neither? Decades of research have found support for both of these hypotheses, but conflicting results prevent consensus (Allen et al., 2009; Mayer & Trezek, 2014; Paul et al., 2009). A meta-analysis of studies investigating the use of phonological information by deaf readers found that only half of the studies showed evidence for the use of phonology by deaf individuals during reading (Mayberry et al., 2011). Further, the meta-analysis indicated that on average, phonological skills predicted 11% of the variance in reading proficiency in deaf individuals, while overall language ability (i.e. skill in a spoken or signed language, independent of reading) predicted 35% of the variance in reading proficiency. Additionally, though there have been some reports of deaf readers experiencing phonological interference effects in the same way that hearing readers do (i.e., slower performance on reading tasks in which phonological similarity between words is a distractor), a 1983 study by Treiman and Hirsh-Pasek found that high-achieving deaf readers did not experience these phonological interference effects. Rather, high-achieving deaf readers showed interference while reading English words that were similar in ASL handshape, suggesting that they may have been relying on their knowledge of ASL while reading English (Treiman & Hirsh-Pasek, 1983).

By contrast, though some studies find that sign language skill is the best predictor of reading proficiency in deaf individuals, other studies fail to find that association (Mayer & Akamatsu, 2011; Moores & Sweet, 1990). Additionally, because most deaf children who grow up in a sign language-rich environment have deaf parents, factors such as an earlier recognition and quicker acceptance of the hearing loss may confound the conclusion that it is a rich early exposure to sign language that aids those deaf children in reading better (Goldin-Meadow & Mayberry, 2001).

1.2. How does anyone become a proficient reader? The debate around how deaf children learn to read is often presented in terms of asking: do deaf children read in the same ways that hearing children do, albeit with reduced direct access to phonological information, or do they read in different ways (Hanson, 1989; Mayer & Trezek, 2014; Perfetti & Sandak, 2000; Wang et al., 2008; Wang & Williams, 2014)? However, research has shown that even hearing individuals attain reading proficiency in a variety of ways – there is not just one pathway to becoming a proficient reader. Many sub-component processes contribute to successful reading comprehension, and these underlying abilities may contribute differentially in different individuals to produce similar reading success. For example, in hearing individuals, differences in reading ability level are associated with both word recognition skill and higher-level cognitive processes such as working memory, a skill that is directly related to syntactic and semantic processing (King & Just, 1991; Perfetti, 1985; Prat & Just, 2011) Better word recognition skill is associated with different reading-related cognitive processes (as assessed by fMRI) than the cognitive processes associated with improvement in working memory (Prat & Just, 2011).

Given that hearing readers can attain reading proficiency in different ways, it is possible that deaf and hearing individuals read successfully using different routes, and also that deaf readers from different language backgrounds may read successfully by different mechanisms. The two main hypotheses about what leads to deaf literacy challenges (lack of phonological knowledge or lack of language proficiency) have dominated the literature for decades. As has been seen, there is evidence and counterevidence for both hypotheses. Some deaf children raised in oral environments with a focus on phonology become proficient readers; some do not. Some deaf children raised in an ASL-rich environment become proficient readers, but not all do. Rather than thinking about the two competing hypotheses as simply all-or-none options, perhaps we need to think more broadly about what proficient deaf readers are doing when they read, and how that relates to the language background they grew up in. Reading education for deaf children could be improved by more thoroughly evaluating the validity of the two aforementioned hypotheses, but not being constrained by them, recognizing that each or neither may hold true in certain circumstances.

1.3. How do we study language processing? Prior research into the mechanisms by which deaf individuals read has primarily used behavioral tasks, such as reaction time measures and standardized reading tests. While much has been learned from this work, the field lacks information on how the brains of deaf readers process language in real time, not mediated by any behavioral task. Such evidence would help identify the neurocognitive mechanisms by which deaf people read successfully, and indicate how these mechanisms differ based on a deaf person’s language background, as well as determine if deaf and hearing individuals read proficiently using the same mechanisms.

1.4. Event-related potentials (ERPs). Event-related potentials (ERPs), recorded while a person reads, are a direct millisecond-by-millisecond record of the brain’s electrical activity, and provide a powerful way to better understand how exactly deaf readers read. ERPs reflect postsynaptic electrical activity of pyramidal neurons in the cortex of the brain, time-locked to a stimulus, and are measured from scalp electrodes. In literacy research, ERPs are recorded while a person reads, and the stimuli are words. ERPs provide information about the brain’s response on a millisecond timescale, which is crucial when studying the rapid processing of language during reading. Two facts about ERPs make them especially well-suited for studying reading.

1.4.1. Language-specific ERP responses. First, ERPs respond to specific aspects of language, including phonology, orthography, meaning, and grammar (Osterhout, McLaughlin, Kim, Greenwald, & Inoue, 2004). Grammatical errors in sentences typically elicit a positive-going peak starting around 500-600ms in an ERP wave (a P600 component), while semantic (meaning) errors in sentences elicit a negative-going component peaking around 400ms (an N400) (Kaan, Harris, Gibson, & Holcomb, 2000; Kutas & Federmeier, 2000; Kutas & Hillyard, 1980, 1984; Osterhout, Holcomb, & Swinney, 1994; Osterhout & Holcomb, 1992; Osterhout & Nicol, 1999). When a word in a sentence is anomalous in both grammar and semantics (i.e., *"The cat will baking the food."*), both an N400 and a P600 are elicited in a nearly additive fashion (Osterhout & Nicol, 1999). The primary exception to this generalization is that phenomena involving the syntax/semantics interface (e.g., assignment of thematic roles) elicit N400s, P600s, or a mixture of the two effects (Kim & Osterhout, 2005; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; van de Meerendonk, Kolk, Vissers, & Chwilla, 2010; van Herten, Chwilla, & Kolk, 2006). If someone is comprehending a natural language that they learned as their native language from a very young age, N400s and P600s are generally elicited as described above, regardless of the language or language modality. Indeed, deaf native users of ASL show N400s when viewing ASL sentences with semantic violations and P600s when viewing ASL sentences with violations of ASL grammar (Capek et al., 2009). Similar results have been found in deaf native signers of German Sign Language (Hänel-Faulhaber et al., 2014).

In addition to sentence-level ERP responses, N400 amplitude is also influenced by word priming. When two words are presented in succession, the N400 to the second word is smaller if the two words are related in some way - phonologically (“cone” “own”), orthographically (“gown” “own”), or semantically (“king” “queen”) (Coch, Grossi, Coffey-Corina, Holcomb, & Neville, 2002; Grossi, Coch, Coffey-Corina, Holcomb, & Neville, 2001; Khateb et al., 2007; Kramer & Donchin, 1987; Rugg, 1984; Weber-Fox, Spencer, Cuadrado, & Smith, 2003).

1.4.2. How does language exposure affect ERP responses? The second aspect of ERPs that makes them so well suited for studying reading and reading proficiency is that prior work has shown that the size of an ERP response varies as a function of a person’s skill and exposure to a language. These results have been seen in people who are in the process of acquiring a second language (L2), people who have become proficient in an L2, and even people comprehending in their first language (L1) who differ in proficiency.

Classically, ERPs have been analyzed by recording data from a group of relatively homogenous participants and averaging the data from all participants together. This provides an overall view of how a particular aspect of language is processed by a particular population, and is generally beneficial for increasing the signal-to-noise ratio. Recently, however, ERP researchers have begun to appreciate that groups generally considered “homogeneous” in fact have quite a bit of variability within them – and importantly, that the variation is associated with systematic differences in ERP responses. The growing body of research on individual differences in ERP-indexed language shows that while variation amongst participants was once considered “noise”, this “noise” is often directly related to proficiency or other differences between the participants (see below for specifics). Rather than being considered a nuisance, these individual differences can be used to learn more about how variations across users of a language are related to how language is processed in the brain.

1.4.2.1. *Second language (L2) processing during classroom learning.* Some ERP research studies have tracked college students as they are learning an L2 in a classroom setting. In some early-stage L2 learners, grammatical errors in sentences elicit an N400 effect rather than a typical grammatical response (a P600). Other learners at that same stage of learning do show a P600 effect in response to grammatical violations. As the learners spend more time studying a language and increase in proficiency, more and more of them display a P600 in response to grammatical errors (McLaughlin et al., 2010; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013).

1.4.2.2. *Second language (L2) processing outside a classroom setting.* In general, higher-proficiency groups of L2 users, or those exposed to their L2 earlier in life, show more native-like ERP responses than do lower-proficiency or later-exposed L2 users.

Rossi and colleagues (2006) found that in response to grammatical errors in German sentences, a higher proficiency group of German L2 users had P600s more similar to a group of native speakers of German than did a group of lower proficiency German L2 speakers (Rossi, Gugler, Friederici, & Hahne, 2006). Weber-Fox and Neville (1996), working with a group of native Chinese speakers with English as an L2, found that individuals exposed to English later in life showed less robust P600s in response to English grammatical violations (Weber-Fox & Neville, 1996).

Though both of these prior studies analyzed ERP data in groups of people, correlation and regression analyses have also shown reliable individual differences in ERP responses. Tanner and colleagues (2014), working with L1 Spanish – L2 English bilingual adults, found via multiple regression that proficiency and language exposure predicted both the size and type of responses to grammatical violations in English sentences. Specifically, they found that better

English proficiency predicted larger magnitude responses to grammatical violations, independent of whether that response was an N400 and a P600. They also found that earlier immersion in an English-speaking environment and great motivation to sound like a native speaker of English predicted a more P600-like response (rather than an N400-like response) to the grammatical violations (Tanner, Inoue, & Osterhout, 2014). This study also indicated that while we generally had thought that as individuals become more proficient in an L2, they progress from showing N400s to showing P600s in response to grammatical violations in sentences (McLaughlin et al., 2010; Tanner et al., 2013), even in highly proficient L2 populations, there are still some people who show N400s in response to grammatical violations (Tanner et al., 2014).

In terms of semantic violations in sentences, Newman and colleagues (2012) worked with L1 Spanish - L2 English speakers and looked at how individual differences in ERP responses to semantic violations related to English proficiency. They found that people who were more proficient in English, their L2, had larger N400s in response to semantic violations in English sentences (Newman, Tremblay, Nichols, Neville, & Ullman, 2012). Ojima and colleagues (2011), working with Japanese schoolchildren learning English as an L2, found via regression analysis that both more hours of exposure to English and later exposure in life to English predicted large N400s in response to semantic violations in sentences (Ojima, Matsuba-Kurita, Nakamura, Hoshino, & Hagiwara, 2011).

Individual differences in responses to semantic priming are also seen in L2 users. ERPs of more proficient L2 users show larger priming responses to semantically related words than do ERPs of less proficient L2 users (Kotz & Elston-Güttler, 2004; Phillips, Segalowitz, Brien, & Yamasaki, 2004).

1.4.2.3. *Individual differences in first language (L1) processing.* While one might think all L1 users are "proficient" in their L1, there is a range of language proficiency within L1 populations (Hammill, Brown, Larsen, & Wiederhold, 1994). Though language-related ERP responses are more homogenous within a group of people with the same first language, differences are still seen based on proficiency. These differences are seen both in ERP responses to sentence violations and in word priming paradigms.

While the P600 is the most common response to grammatical violations in sentences, some L1 users produce an N400 to grammatical violations (Osterhout, McLaughlin, Kim, Greewald, & Inoue, 2004; Osterhout, 1997; Tanner & Van Hell, 2014). Additionally, though the N400 is the most common response to semantic violations in sentences, groups of L1 users with higher working memory abilities have shown P600 responses to some semantic violations (Nakano, Saron, & Swaab, 2010).

In terms of what behavioral or subject information is associated with different ERP responses in L1 users, Pakulak and Neville (2010) found that higher L1 proficiency correlated with larger P600s in response to grammatical violations (Pakulak & Neville, 2010). In response to semantic errors in sentences, there is some discrepancy in the literature. One study found that a group of highly proficient L1 participants showed a smaller N400 response to semantic violations than did a group of L1 participants with average proficiency levels (Weber-Fox, Davis, & Cuadrado, 2003). A separate study found nearly the opposite, that higher L1 proficiency was correlated with a larger N400 response to semantic violations (Newman et al., 2012).

There are also systematic individual differences in ERP word priming responses in an L1. Semantic priming in an L1 appears to be related to proficiency; higher L1 proficiency is associated with stronger N400 semantic priming effects in that language (Balass, Nelson, & Perfetti, 2010; Connolly, Byrne, & Dywan, 1995; Friedrich & Friederici, 2006; Henderson, Baseler, Clarke, Watson, & Snowling, 2011; Landi & Perfetti, 2007; Perfetti, Wlotko, & Hart, 2005). Children with dyslexia who have poor phonological coding skills show reduced or altered ERP priming effects to words related in phonology (Jednorog, Marchewka, Tacikowski, & Grabowska, 2010; McPherson, Ackerman, Holcomb, & Dykman, 1998).

1.4.3. ERP reading research in deaf readers. Aside from one study investigating ASL comprehension in deaf participants, all of the aforementioned research has been performed in hearing individuals. What ERP responses do we see when deaf individuals are reading a written language? Though the literature on this topic is quite limited, there is useful information to be taken from it.

Early work by Helen Neville and colleagues looked at ERP responses to individual written words in deaf and hearing adults. In a 1984 study motivated by investigating hemispheric differences in language processing, Neville, Kutas, and Schmidt found that when shown English nouns, deaf readers had different hemispheric patterns of ERP responses than did hearing readers (Neville, Kutas, & Schmidt, 1984). In a later study by Neville, Mills, and Lawson (1992), deaf native signers and English L1 hearing adults read two different types of English words as part of sentences: open class words, which mainly provide semantic information (e.g. nouns, adjectives, most verbs), and closed class words, which mainly provide grammatical information (e.g. prepositions, articles, pronouns). Some of the sentences had semantic violations in them, and some did not. They found deaf participants' ERP responses to open class words were similar to hearing participants' responses, but the deaf participants' responses to closed class words were qualitatively different than what was seen in the hearing participants. They also found that both deaf and hearing participants showed larger N400s in response to semantic violations in sentences. This early work led the researchers to conclude that semantic language processing was relatively similar between the deaf and hearing participants, but that grammatical language processing was more affected by the different language backgrounds (Neville, Mills, & Lawson, 1992).

A more recent study by MacSweeney, Goswami, and Neville (2013) investigated ERP responses to phonological priming in severely and profoundly deaf adults. They found that deaf participants who were able to perform at above-chance levels on a phonological judgement task showed a reduction in N400 amplitude in response to words related only in phonology – a typical ERP priming response (MacSweeney, Goswami, & Neville, 2013). Results from deaf participants who did not perform at above-chance levels on the phonological judgement task were not presented.

One research group has recently used ERPs to investigate how deaf participants respond when reading sentences with semantic and grammatical violations (Skotara, Kügow, Salden, Hänel-Faulhaber, & Röder, 2011; Skotara, Salden, Kügow, Hänel-Faulhaber, & Röder, 2012). In their first study, three groups of participants read German sentences with semantic and grammatical errors: 1) hearing participants with German as an L1, 2) hearing participants with German as an L2, and 3) profoundly congenitally deaf native signers of German Sign Language (DGS) with written German as an L2. In response to semantic violations in sentences, all three groups showed an N400 and a later positivity, with no differences in N400 amplitude between any of the groups. In response to grammatical violations, while all three grows showed a P600, the hearing participants with German as an L2 had smaller P600s than the hearing participants with German as an L1. There was no difference in P600 amplitude between the hearing participants with German as an L1 and the deaf native signers. The researchers concluded that even though written German was learned mainly via a visual route by the deaf native signers, the patterns of neural language processing were similar to what was seen in hearing readers with German as an L1, and that learning a sign language as a native language "results in comparable functional organization of the L2" (Skotara et al., 2011).

As has been previously discussed, deaf native signers are thought to have an advantage when learning to read, because they have attained proficiency in a language, as compared to many deaf children who do not grow up with access to a language that they can learn naturally, and are not proficient in any language when they begin to learn to read. Skotara and colleagues (2012) investigated this question in a second study, in which participants were 1) hearing participants with German as an L1, 2) profoundly congenitally deaf native signers of German Sign Language (DGS) with written German as an L2, who had deaf parents, and 3) profoundly congenitally deaf participants with hearing parents who did not learn either DGS or written German until school. This third group essentially had both German Sign Language and written German as L2s; they grew up in a language-impoverished environment, and did not learn sign language naturally like the native signers did. As in the 2011 study, all participants read German sentences with semantic and grammatical violations. In response to semantic violations in sentences, all three groups showed an N400 and a later positivity, with no differences in N400 amplitude between any of the groups. In response to grammatical violations, while all three grows showed a P600, the deaf non-native signers had smaller P600s than the hearing participants. There was no difference in P600 amplitude between the hearing participants and the deaf native signers (Skotara et al., 2012). These results show striking similarity to the first study, in which the hearing participants with German as an L2 had smaller P600s than the hearing participants with German as an L1. Skotara and colleagues concluded that semantic information processing was not particularly affected by early L1 experience (or lack thereof), whereas grammatical language process did seem to be.

The research done by Skotara and colleagues (2011, 2012) gives us great insight into how groups of hearing L1, hearing L2, deaf native signers and non-native signing deaf individuals process language in similar and different ways. Overall, these studies suggest that all groups processed information about semantics in similar ways, but that hearing L2 and non-native signing deaf individuals show differences in how their brains process grammatical information, as compared to hearing L1 and deaf native signers.

However, these studies have limitations that prevent even more information from being learned. As has been described earlier, even people with similar language backgrounds have systematic differences in their ERP responses. In both of these studies, there was no analysis of individuals' ERP responses as compared to their reading comprehension skill. The 2012 study involving native and non-native signing deaf individuals did measure German reading and grammar proficiency in all deaf participants, and found that for the participants whose data was included in the ERP analysis (see below for details), there was no difference in any of the proficiency measures between the groups of deaf native and non-native signers. It is notable that even though the two groups of deaf participants were at similar overall proficiency levels, the non-native signing deaf individuals had smaller P600s than the hearing L1 group, while the native signers did not. German reading proficiency was not measured in the hearing L1 participants, so the proficiency levels of the two groups of deaf participants cannot be compared with the proficiency of the hearing group (Skotara et al., 2012).

Additionally, in both studies, only participants who judged sentence acceptability at above-change levels were included in the ERP analysis, and only trials in which participants correctly judged sentence acceptability were included. While this may serve as an attempt to only include participants with higher reading comprehension skill when they correctly judged sentence acceptability, prior ERP research has shown that even when individuals may not be able to consciously report linguistic knowledge, ERP responses can still indicate recognition of linguistic information. This is shown most readily in a study by McLaughlin, Osterhout, and Kim (2004), in which college students who were just starting to learn French as an L2 were shown individual French words and pseudowords. Pseudowords are combinations of letters that follow a language's phonological rules but are not actually words, and are known to elicit larger N400s than real words. After only 14 hours of classroom instruction, while the French learners could not reliably distinguish between real and non-real French words in a behavioral judgement task, their N400s were significantly larger in response to the French pseudowords than the real French words (McLaughlin, Osterhout, & Kim, 2004). This showed that ERPs can reflect linguistic knowledge that participants may not consciously be able to report. Thus, important information is potentially lost by excluding participants who performed at lower levels of accuracy as well as trials in which incorrect sentence acceptability judgements were given.

In summary, while the research by Skotara and colleagues is important for showing that hearing L1 participants and deaf native signers show similar patterns of grammatical language processing, there is still much to be learned about the relationship between reading comprehension skill and ERP responses to language in deaf readers.

1.5. Goals of this study. The goal of this research study was to use real-time measures of language processing (ERPs) to understand how some deaf individuals read more proficiently than others. Two specific questions we aimed to answer were a) do deaf and hearing individuals read proficiently using the same language processing mechanisms?, and 2) do deaf individuals from different language backgrounds read proficiently using the same language processing mechanisms? To get as full a picture as possible about what aspects of language processing are associated with more proficient reading, we used two different types of ERP stimuli: sentence stimuli and word-pair stimuli. The sentence stimuli, with errors of grammar and semantics, provided information on how participants processed higher-level language concepts of meaning and grammar. The word-pair stimuli, with words related in phonology, orthography, and semantics, provided information on how participants processed specific information about words while reading.

In addition to typical grand mean ERP analyses, we compared the magnitude of participants' ERP responses to their standardized reading skill using correlation and regression analysis. If deaf and hearing participants read proficiently using similar strategies or neural mechanisms, we expected to see similar relationships between higher reading skill and the magnitude of our measured ERP effects in both groups. If we saw different relationships between reading skill and ERP effect magnitudes between deaf and hearing participants, that would be an indication that the two groups were reading proficiently using different strategies or neural processing mechanisms. Within the group of deaf participants, we also examined the relationships between these variables when language background was taken into account, so that we could determine if deaf individuals from different language backgrounds might be reading proficiently in different ways.

In addition to giving us a thorough picture of what patterns of language processing are associated with proficient reading, by comparing the relationship between ERP effects of phonological priming and reading proficiency, we were able to specifically examine the hypotheses that phonological knowledge is important for deaf individuals to become successful readers. In terms of evaluating the hypothesis that early language proficiency is important for deaf individuals to become successful readers, we had very few deaf participants who were native signers in this study, and thus were unable to make a complete analysis of this question. However, our sample of primarily non-native signing deaf participants gave us the ability to compare the neural mechanisms of reading processing between non-native signing deaf reader and hearing readers. Given that 90% of deaf children are born to hearing parents, and thus that the large majority of deaf children do not grow up with a sign language as their native language, understanding how non-native signing deaf individuals read proficiently is of the utmost importance.

Chaper 2. Method

2.1. Participants.

2.1.1. General characteristics. Participants were 42 deaf and 42 hearing adults. The number of participants needed was determined via power analysis, described later in this manuscript (see section "Individual differences analyses"). All deaf participants were severely or profoundly deaf (hearing loss of 71 dB or greater, self-reported), except for one participant with profound (95 dB) hearing loss in the left ear and moderate (65 dB) hearing loss in the right ear. All participants lost their hearing by the age of two. 33 of the 42 deaf participants reported being deaf from birth. Three of the remaining participants reported that it was likely they were deaf from birth but had not been diagnosed until later (still by age two). The final 6 participants reported clear causes of deafness that occurred after birth but before age two. Deaf participants came from a wide variety of language backgrounds, and were asked in detail about their spoken and manual/signed language exposure and use throughout their life. This information is reported in a later section. All deaf participants reported having worn hearing aids in one or both ears at some point in life; 22 participants reported still wearing hearing aids, 5 participants reported only wearing them occasionally or in specific circumstances, and 15 participants reported no longer wearing them. One participant, age 28.5 years, had a unilateral cochlear implant, but it was implanted late in life (at age 25.75 years) and the participant reported rarely using it. Other than that, individuals with cochlear implants did not take part in this study.

Each hearing participant was matched in age and gender to a deaf participant, in order to reduce variability between the two groups. The first language of all hearing participants was English, and English was the only language that had been used in their homes while they were growing up. In each group of participants there were 27 women and 15 men. The average age of deaf participants was 38.63 years and the average age of hearing participants was 38.65 years; there was no significant difference in the ages of the two groups ($t = -0.011$, $p = 0.991$). More details about age and other participant characteristics can be found in Table 1.

Table 1. Participant characteristics and background.

	Deaf			**Hearing**			Groups sig. different?
Measure	Mean	St. Dev.	Range	Mean	St. Dev.	Range	
Age at testing (years)	38.6	11.2	19.6-62.7	38.7	11.2	19.3-63.7	No (p=0.991)
Years of education	16.5	2.1	13-20	17.3	2.5	13-25	No (p=0.111)
NTID Speechreading score (max 100)	59.3	21.4	13.5-89.5	43.7	13.6	15-73.5	t=3.984, p<0.001
WRMT Reading Comprehension (max 124)	82.8	18.8	40-115	101.7	12.4	46*-116	t=-5.449, p<0.001
Deaf participant language background	1=all spoken; 7=all manual/signed						
Growing up language use	4.0	2.1	1-7				
Current language use	5.4	1.9	1-7				
Self-rated ASL proficiency							
1-7 scale (7=fluent, 1=no proficiency)	6.0	1.2	2-7				
14-point descriptive scale							
Comprehension (max 7)	6.2	1.4	1-7				
Production (max 7)	6.2	1.4	1-7				
Total (max 14)	12.4	2.7	2-14				

*Effective range 80-116; second lowest score was 80.

All participants (deaf and hearing) had normal or corrected-to-normal vision, except for one deaf participant with reduced peripheral vision due to Usher syndrome. The deaf participant with Usher syndrome did not have any difficulty in completing any of the study procedures. No participants had any history of significant head injury or epilepsy. While most ERP studies of language restrict participants to only right-handed participants, left-handed participants were allowed in this study so that as many deaf individuals as possible could participate. Two of the deaf participants and seven of the hearing participants were left-handed, as assessed by an abridged version of the Edinburgh Handedness Inventory (Oldfield, 1971). One deaf participant and one hearing participant were ambidextrous. Handedness was not matched across deaf and hearing participants.

All participants filled out a detailed life history questionnaire that asked about many variables, including language background (see below). Other than questions about language background and hearing loss, deaf and hearing participants answered the same questions about age, gender, race/ethnicity, history of vision problems, socioeconomic status, and years of education completed. Years of education completed was counted starting at kindergarten through as much education (including postgraduate degrees) as had been completed by each participant. The deaf and hearing groups did not differ significantly in the number of years of education completed ($t = -1.611$, $p = 0.111$; Table 1).

2.1.2. Language background. Deaf participants completed quantitative and qualitative measures of exposure to spoken versus manual/signed language at different points in life. On a 1-to-7 scale, where 1=all oral communication, 7=all manual/signed communication, and 4=equal use of both, participants were asked about their method(s) of communication at the following points in their life: a) during primary and secondary education, b) overall while they were growing up (incorporating language use both in school and in the home), and c) at the current point in time. Importantly, a '7' on this scale did not distinguish between the use of American Sign Language (ASL) and manually coded forms of English (i.e., Signed Exact English, SEE). Participants also wrote descriptions of their language at each of these points in time, which served two purposes. First, it allowed for the researcher to confirm that the participants' ratings on the 1-to-7 scales matched what they described – and if they did not, to ask for clarification from the participant. Second, it allowed us to distinguish between participants who grew up using and being exposed to ASL versus those who grew up using and being exposed to forms of Manually Coded English.

The language background of the deaf participants was extremely diverse. On all of the 1-to-7 language scales, responses ranged from 1 to 7. The average response on the scale about language use while growing up (referencing language use both in school and in the home) was a 4.0, indicating that there was a nearly equal mixture of participants who grew up in more a spoken language environment versus a more manual/signed language environment. Though many participants reported using ASL, SEE, and/or PSE at some point while growing up (at home and/or at school), only four participants had deaf parents or other family members who communicated with them in fluent ASL from birth or a young age. Of those four participants, two participants had biological parents who were deaf, one participant was adopted at 18 months by deaf parents, and one participant had two older sisters who were deaf. Thus, while the deaf participants in this study came from a wide range of spoken versus manual/signed language environments, few participants were truly native ASL signers.

The average response on the scale for current language use was 5.4, indicating that at the current point in time, there was greater use of manual/signed communication than spoken communication amongst our participants. More information on the results of these language measures can be found in Table 1.

Deaf participants were also asked to self-rate their ASL proficiency. All participants rated themselves on a 1-to-7 scale, where 1 indicated no proficiency in ASL and 7 indicated being “fluent and completely comfortable communicating in ASL”. The scale instructions explicitly said to consider only ASL proficiency, not proficiency in PSE or SEE. While responses ranged from 2 to 7 on the scale, the average was a response of 6.0; the large majority of the participants rated themselves at the highest levels of ASL proficiency (Table 1). This skew was seen after an initial 16 deaf participants took part in the study. In an effort to use a better self-report measure, both to better distinguish between ASL proficiency levels and to improve the validity of the self-rated ASL proficiency measure, a more detailed ASL self-report scale was used after the initial 16 participants took part in the study. In this scale, seven statements about ASL comprehension ability and seven statements about ASL production were given, with the statements increasing in difficulty from “You can understand several words in ASL” (statement 1 of 7, easiest) through “You understand ASL well enough to follow a casual conversation with a native signer” (statement 4 of 7, middle difficulty) to “ASL is your dominant (strongest) language or you understand ASL as well as your strongest language” (statement 7 of 7, hardest difficulty) (Stocco & Prat, 2014). Statements about production were similar. Participants marked if each statement was true or false of their ASL comprehension and production ability. The number of true statements was totaled in each of the comprehension and production subsections, and a total score was also calculated as the sum of both of those subsections. The initial 16 participants who took part before we introduced this more detailed proficiency scale were re-contacted to fill out the new measure, but we were unable to obtain responses from 5 participants. Despite this being a much more detailed self-report scale of ASL proficiency, the results still showed that the deaf participants rated themselves as highly proficient in ASL. The average score from both the comprehension and production sections was 6.2 (max 7), making the total average score 12.4 out of 14 possible (Table 1).

Hearing participants' questions about language background were used to ensure that English was their first language and that they had not been exposed to other languages in the home early in life. Most hearing participants had been exposed to other languages while in school; proficiency in those later-learned languages varied.

2.1.3. Standardized reading comprehension. Since the goal of this research was to investigate which ERP responses predict better reading skill, a measure of standardized reading was required. We measured additional variables that are known to be correlated with reading skill, so that the variability in reading skill associated with other factors could be accounted for during analysis.

Standardized reading comprehension was measured using the Word and Passage Comprehension sections of the Woodcock Reading Mastery Tests, Third Edition (WRMT-III) (Woodcock, 2011). These two sections make up the Reading Comprehension portion of the WRMT-III. The WRMT-III was used because it can be administered to participants of all reading levels. Because great variability was expected in the reading level of deaf participants, it was important to have a single test that could be used across a broad range of reading levels. Though the test typically requires verbal answers, deaf participants were instructed to respond in their preferred method of communication; this test is often administered in this way to deaf participants (Easterbrooks & Huston, 2008; Kroese, Lotz, Puffer, & Osberger, 1986; L. Spencer, Barker, & Tomblin, 2003). In the case of participants who sign their answers, there could have been concern that an ASL sign might represent multiple English words. We worked closely with our interpreters to ensure that when the specific English word being signed was unclear, the participant was asked fingerspell the English word they were intending.

The maximum total raw score possible on the Reading Comprehension section of the WRMT-III was 124. The average score of deaf participants (mean = 82.8, standard deviation = 18.8) was significantly lower than the average score of hearing participants (mean = 101.7, standard deviation = 12.4; $t = -5.449$, $p < 0.001$; Table 1). However, the highest scoring deaf participants scored as high as the highest scoring hearing participants. Deaf participants ranged in score from 40-115. Hearing participants effectively ranged in score from 80-116; a single hearing participant had a score of 46 (Figure 1). 24 of the 42 deaf participants had scores within the effective range of hearing participants. Thus, while the deaf participants had a much wider range of scores, the highest scoring participants in both groups were at the same reading level.

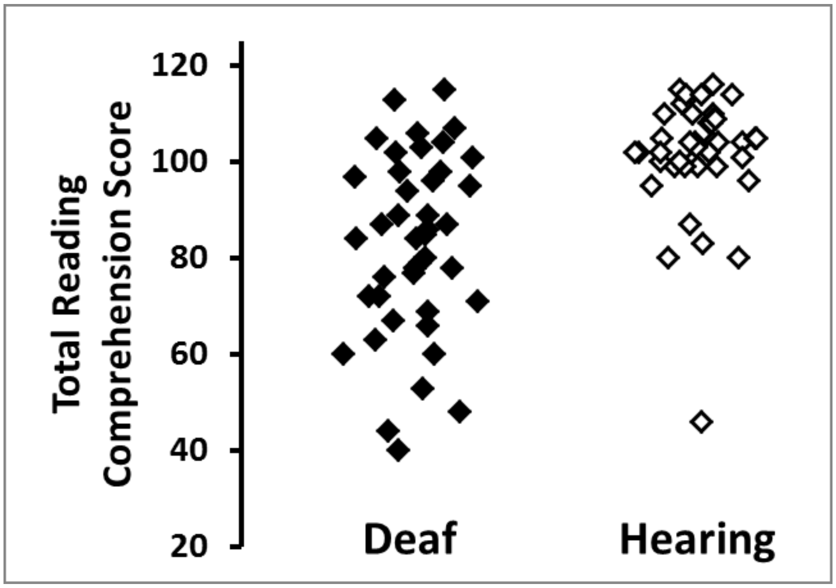


Figure 1. Distribution of standardized reading comprehension scores for deaf and hearing participants.

There was no significant correlation between the language background while growing up of the deaf participants and their standardized reading comprehension scores (r = -0.168, p = 0.288; Figure 2).

2.1.4. Speechreading skill. Standardized speechreading skill was measured in all participants, due to the fact that speechreading is highly correlated with reading proficiency in deaf individuals (Kyle & Harris, 2010; Mohammed, Campbell, MacSweeney, Barry, & Coleman, 2006). Speechreading skill was measured with the National Technical Institute for the Deaf's

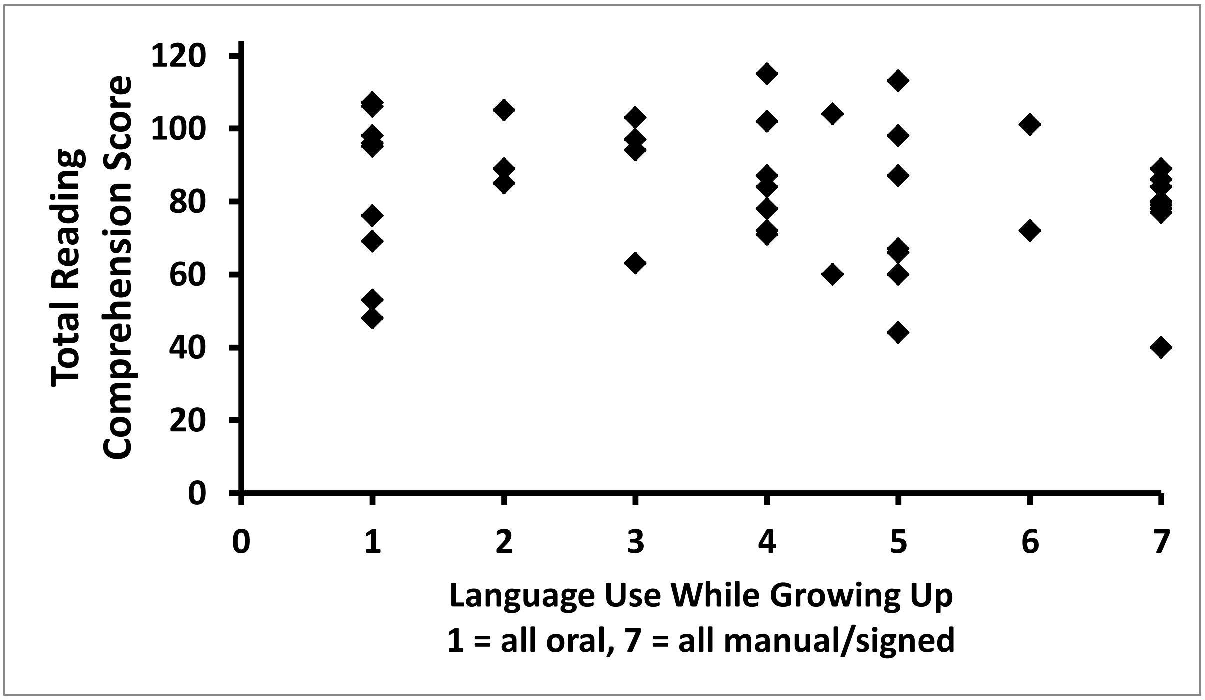


Figure 2. Relationship between deaf participants' language use while growing up and reading comprehension score.

"Speechreading: CID Everyday Sentences Test" (Sims, 2009). In this test, participants watch a video, without sound, of a speaker saying 10 sentences. After each sentence, participants write what they believe the speaker said. The test is scored based on the number of correct words documented. The test comes with 10 lists of 10 sentences each; participants were evaluated on the sentences from List 1 and List 6, and the average of the scores from those two lists was taken. For every possible raw score with each list, there was also a corrected score used to normalize slight difficulty differences between lists. Lists 1 and 6 were chosen because the relationship between their raw and corrected scores was similar. Corrected scores are reported here.

The maximum average speechreading test score was 100. Overall, the group of deaf participants performed significantly better on the speechreading test than the hearing participants (deaf: mean = 59.3, standard deviation = 21.4; hearing: mean = 43.7, standard deviation = 13.6; $t = 3.984$, $p < 0.001$; Table 1). However, there was large variability within each group. Deaf participants' scores ranged from 13.5-89.5; hearing participants' scores ranged from 15-73.5. A visualization of the ranges of speechreading scores between the two groups can be seen in Figure 3.

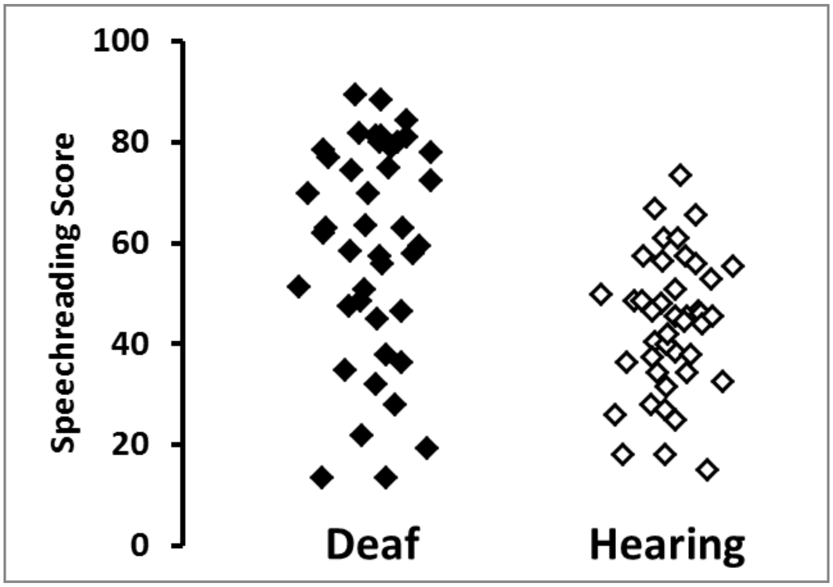


Figure 3. Distribution of standardized speechreading scores for deaf and hearing participants.

As was expected, in deaf participants, higher speechreading test score was strongly correlated with higher standardized reading comprehension score ($r = 0.702$, $p < 0.001$). For hearing participants, there was no significant relationship between speechreading test score and standardized reading comprehension score ($r = 0.173$, $p = 0.273$).

2.1.5. Frequency of reading habits. Finally, frequency of reading habits was measured in multiple ways. Participants completed a questionnaire that asked how often they currently read for pleasure, for work or school, and how often they read while growing up. The questionnaire also asked what types of things they read at each point in time. Participants also completed the Author Recognition Test (ART), which provides an unbiased way to quantify amount of exposure to printed material (Stanovich & West, 1989). These reading frequency measures were not introduced until after the first 16 deaf and first 15 hearing participants had already taken part in the study. Those initial participants were re-contacted to fill out the new measures, but we were unable to obtain responses from 6 deaf participants and 3 hearing participants. Thus, these reading frequency measurements were not used in the final analysis.

2.2. Materials. Two types of ERP stimuli were used in this experiment: sentences for a sentence violation paradigm, and word pairs for a priming paradigm.

2.2.1. Sentence stimuli. Sentence stimuli were 120 sentence quadruplets in a full crossed 2 (semantic correctness) by 2 (grammaticality) design. Sentences were either grammatically correct or contained a subject-verb agreement violation, and were also either semantically well-formed or contained a semantic anomaly. This resulted in four types/conditions of sentences: 1) well-formed sentences, 2) sentences with a grammatical (syntactic) violation alone, 3) sentences with a semantic (meaning) violation alone, and 4) sentences with a double violation – a combination grammatical and semantic violation. All violations, semantic and/or syntactic, occurred on the critical word in the sentence. Critical words were either verbs in their base/uninflected form (e.g., belong) or in their third person singular present tense form (e.g., belongs). Each sentence condition (well-formed, grammatical violation, semantic violation, double violation) had two versions – one with the critical word in the base form and one in the –s form (see Table 2 for an example). The stimuli were

Table 2. Example sentence stimuli.

Condition	Sentences
Well-formed	The huge house still belongs to my aunt. The huge houses still belong to my aunt.
Grammatical violation	The huge houses still belongs to my aunt. The huge house still belong to my aunt.
Semantic violation	The huge house still listens to my aunt. The huge houses still listen to my aunt.
Double grammatical & semantic violation	The huge houses still listens to my aunt. The huse house still listen to my aunt.

Note: The critical word for ERP averaging is underlined.

designed in this way so that the singular/plural status of the subject of the sentence (the noun preceding the critical verb) could not be used to predict whether or not the sentence would contain an error. Thus, while there were only four different conditions of sentences (well-formed, grammatical violation, semantic violation, double violation), there were eight versions of each sentence (Table 2).

Sentences that were well-formed or contained grammatical violations used a set of 120 unique verbs (in both their base and –s forms), and sentences that contained a semantic or double violation used a different set of 120 unique verbs. The two sets of verbs were chosen so that the average written word-form log frequency (provided in the CELEX2 database (Baayen, Piepenbrock, & Gulikers, 1995)) of the verbs in the two sets was not significantly different (well-formed/grammatical verbs average frequency = 0.59, semantic/double verbs average frequency = 0.56, $t = 0.441$, $p = 0.659$). These calculations took into account the average frequency of both the base and –s forms of the verbs in each set.

The eight versions of each sentence were distributed across eight experimental lists, such that each participant only saw one version of each sentence. There were 15 sentences from each of the eight versions in each list, and thus there were 30 sentences per condition in each list. Each list contained an additional 60 filler sentences, all of which were grammatically correct. In total, each list contained 180 sentences. The sentence order in each list was randomized, and lists were divided into 3 blocks of 60 sentences each. Participants were pseudorandomly assigned one of the sentence lists.

2.2.2. Word-pair stimuli. Word-pair stimuli were words related and unrelated in meaning (semantically related) as well as words related and unrelated in phonology (sound), orthography (spelling), or both phonology and orthography. The first word in a pair is the ‘prime’, and the second is the ‘target’. The semantic stimuli consisted of sixty target words, each of which had both an unrelated and a related prime (see Table 3 for an example). The two versions of each prime-target pair were distributed across experimental lists, such that each participant saw each target word only once; 30 of the targets with a semantically related prime, and the other 30 targets with a semantically unrelated prime. The targets were split in half such that the average written word-form log frequency (provided in the CELEX2 database (Baayen et al., 1995)) of the first 30 targets and the second 30 targets was not significantly different (first half mean = 1.54, second half mean = 1.47, $t = 0.298$, $p = 0.766$). There were 30 semantically related pairs and 30 semantically unrelated pairs per list.

Table 3. Example word-pair stimuli.

Condition	**Example**
Semantically unrelated	king - queen
Semantically related	window - queen
Phonologically + orthographically unrelated	raid - pear
Orthographically related	dear - pear
Phonologically related	lair - pear
Phonologically + orthographically related	wear - pear

Note: The underlined word was used for ERP averaging

The phonologic/orthographic stimuli consisted of 120 additional target words, each of which had four associated primes: 1) an unrelated prime, 2) a prime related only in phonology, 3) a prime related only in orthography, and 4) a prime related in both phonology and orthography (see Table 3 for an example). As with the semantic primes, the four versions of each prime-target pair were distributed across experimental lists, such that each participant only saw each target word once; 30 of the targets with each of the 4 types of primes. The targets were split into four groups such that the average written word-form log frequency (provided in the CELEX2 database (Baayen et al., 1995)) of each of the four sets of 30 targets was not significantly different (means of the four sets: 0.89, 1.01, 0.73, 0.99, $F = 0.685$, $p = 0.563$). There were 30 pairs from each of the four conditions in each list.

The phonologic/orthographic word pairs (120 per list) were in the same lists as the semantically related/unrelated word pairs (60 per list). In addition to these 180 experimental word pairs, each list contained 105 additional pairs in which one word was a pseudoword – a set of characters that follows English orthographic rules but is not a real word (e.g., floop). 90 of the pseudowords were primes in a pair, and 15 were targets in a pair; the other word in all those pair was a real English word. Finally, each list also contained 60 pairs of word in which one word was a nonword – a set of characters that did not follow English orthographic rules. 30 of the nonwords were primes in a pair, and 30 were targets in a pair. This led to a total of 345 word pairs per list. 52.17% of the pairs contained only real English words, and 47.83% of the pairs contained one non-real word. The order of word pairs in each list was randomized, and the lists were divided into 3 blocks of 115 pairs each. Participants were pseudorandomly assigned one of the word pair lists.

2.3. Procedure. Participants took part in three sessions, each of which lasted no more than two hours. An interpreter was present for all sessions with deaf participants except for participants who specified that they did not need an interpreter. Four of the 42 deaf participants specified that they did not need an interpreter present for their visits. In the first session, participants completed all background questionnaires and the speechreading and reading comprehension tests. ERPs were recorded during the second and third sessions. Participants viewed the sentence and word-pair stimuli in separate sessions. Half of the participants saw the sentence stimuli in the second session and the word-pair stimuli in the third session; the other half of the participants saw them in the opposite order. Participants were pseudorandomized to determine the order of their ERP sessions.

During ERP recording, participants were seated in a comfortable recliner in front of a CRT monitor. Participants were instructed to relax and minimize movements and eye blinks while silently reading the stimuli in their minds. Specifics of the procedures differed based on which stimuli were being viewed, described below.

2.3.1. Sentences. Each sentence trial consisted of the following events: a blank screen for 1000 ms, followed by a fixation cross, followed by a stimulus sentence presented one word at a time. The fixation cross appeared on the screen for 500 ms followed by a 400 ms interstimulus interval (ISI). Each word of the sentence appeared on the screen for 600 ms followed by a 200 ms ISI. The presentation rate used is slower than the rate typically used in ERP studies with first language users; it is standard procedure for ERP studies of a second language to use a slower presentation rate for all participants, both the first and second language users (Foucart & Frenck-Mestre, 2011, 2012; Tanner et al., 2014, 2013). After the final word of the sentence, there was a 1000 ms blank screen, followed by a "yes/no" prompt. Participants were instructed to give a sentence acceptability judgment at the "yes/no" prompt, where "yes" was the response for sentences that were correct in all ways and "no" was the response for sentences that contained any kind of error. Participants were instructed to make their best guess if they were not sure whether the sentence contained an error or not. The "yes/no" prompt remained on the screen until participants responded "yes" or "no"; as soon as a response was given, presentation of the next sentence began. Participants were pseudorandomly assigned to use either their left or right hand for the "yes" response.

2.3.2. Word pairs. Each word pair trial consisted of the following events: a blank screen for 1000 ms, followed by a fixation cross, followed by a stimulus word pair presented one word at a time. The fixation cross appeared on the screen for 500 ms followed by a 400 ms ISI. The first word of the pair (the prime) appeared on the screen for 600 ms followed by a 200 ms ISI. The second word of the pair (the target) appeared on the screen for 800 ms. After the target word, there was a 500 ms blank screen, followed by a "yes/no" prompt. Participants were instructed to give a lexical decision judgment about both words in the pair at the "yes/no" prompt, where "yes" was the response when both words in the pair were real English words, and "no" was the response when either word of the pair was not a real English word. Participants were instructed to make their best guess if they were not sure whether the words were real English words or not. The "yes/no" prompt remained on the screen until participants responded "yes" or "no"; as soon as a response was given, presentation of the next pair of words began. Participants used the same hand (left or right) for the "yes" response for both the word pair and sentence stimuli.

2.4. Data acquisition and analysis. Continuous EEG was recorded from 19 tin electrodes attached to an elastic cap (Electro-cap International) in accordance with the 10–20 system (Jasper, 1958). Eye movements and blinks were monitored by two electrodes, one placed beneath the left eye and one placed to the right of the right eye. Electrodes were referenced to an electrode placed over the left mastoid. EEG was also recorded from an electrode placed on the right mastoid to determine if there were experimental effects detectable on the mastoids. No such effects were found. EEG signals were amplified with a bandpass filter of 0.01–40 Hz (-3db cutoff) by an SAI bioamplifier system. ERP waveforms were filtered offline below 30 Hz. Impedances at scalp and mastoid electrodes were held below 5 kΩ and below 15 kΩ at eye electrodes.

Continuous analog-to-digital conversion of the EEG and stimulus trigger codes was performed at a sampling frequency of 200 Hz. ERPs, time-locked to the onset of the critical word in each sentence (underlined in the examples of Table 2) or the target word in each word pair (underlined in the examples of Table 3) were averaged offline for each participant at each electrode site in each condition. Trials characterized by eye blinks, excessive muscle artifact, or amplifier blocking were not included in the averages. ERPs were quantified as mean amplitude within a given time window. All artifact-free trials were included in the ERP analyses. Time windows and analysis specifics differed between word pair and sentence stimuli.

For sentence stimuli, in accordance with previous literature and visual inspection of the data, the following time windows were chosen: 300-500ms (LAN/N400), and 500-900ms (P600), relative to a 100ms prestimulus baseline. Differences between sentence conditions were analyzed using a repeated-measure ANOVA with two levels of semantic correctness (semantically plausible, semantic violation) and two levels of grammaticality (grammatical, ungrammatical).

For word-pair stimuli, in accordance with previous literature and visual inspection of the data, a 300-500ms (N400) time window was used, relative to a 100ms poststimulus baseline. A 100ms poststimulus baseline was used for word pair data because there were moderate differences in the prestimulus baseline, due to differing responses to the primes of the word pairs. For phonologically- and orthographically-related word pairs, differences between the conditions were analyzed using a repeated-measures ANOVA with two levels of phonologic relatedness (phonologically related, phonologically unrelated) and two levels of orthographic relatedness (orthographically related, orthographically unrelated). Semantically-related word pairs were analyzed in a separate ANOVA, with two levels of semantic relatedness (semantically related, semantically related.

All ANOVA analyses, for both sentence and word pair data, included deaf and hearing participants in the same model, using group (deaf or hearing) as a between-subjects factor, unless otherwise specified. For all stimuli, both sentences and word pairs, data from midline (Fz, Cz, Pz), medial (right hemisphere: Fp2, F4, C4, P4, O2; left hemisphere: Fp1, F3, C3, P3, O1), and lateral (right hemisphere: F8, T8, P8; left hemisphere: F7, T7, P7) electrode sites were treated separately in order to identify topographic and hemispheric differences. ANOVAs on midline electrodes included electrode as an additional within-subjects factor (three levels), ANOVAs on medial electrodes included hemisphere (two levels) and electrode pair (five levels) as additional within-subjects factors, and ANOVAs over lateral electrodes included hemisphere (two levels) and electrode pair (three levels) as additional within-subjects factors. The Greenhouse-Geisser correction for inhomogeneity of variance was applied to all repeated measures on ERP data with greater than one degree of freedom in the numerator. In such cases, the corrected *p*-value is reported.

In instances where follow-up contrasts were required, a familywise Bonferroni correction was used based on the number of follow-up contrasts being performed, and the adjusted alpha level is reported.

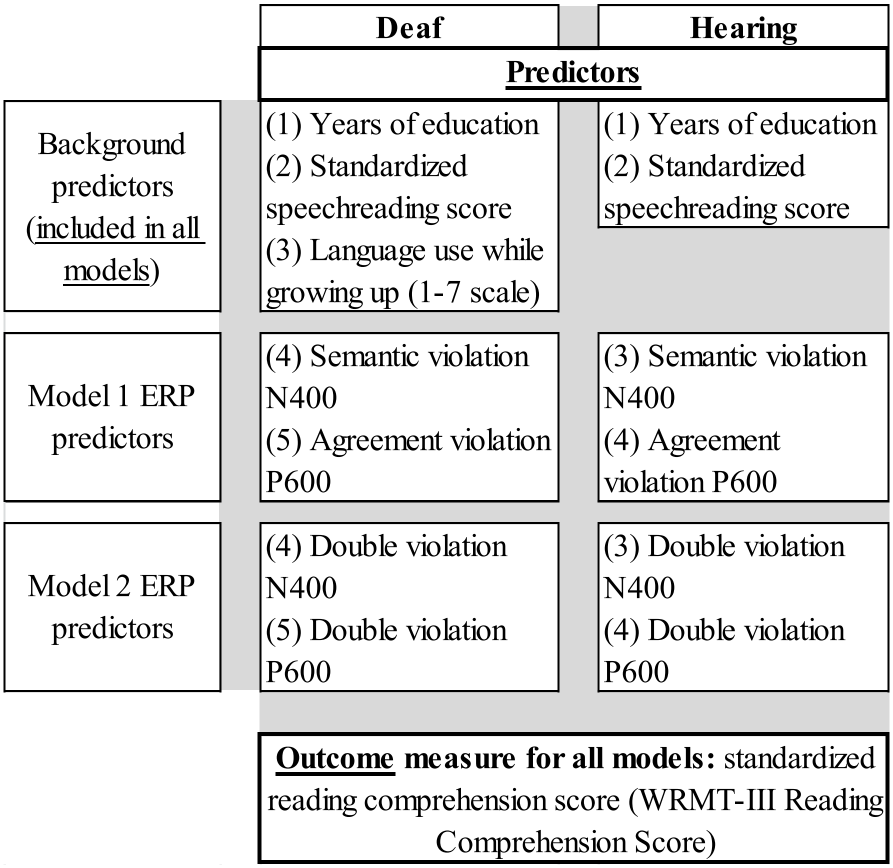
2.4.1. Individual differences analyses. Because the goal of this project was to determine what patterns of ERP responses are associated with better reading skill in deaf and hearing readers, specific individual differences analyses were planned. For all individual differences analyses, the size, or effect magnitude, of a particular ERP responses was calculated. An ERP effect magnitude refers to the “size” of a particular response as compared to the relevant control response. Specifics of these calculations and the associated analyses are described below, and follow similar methods as prior work (Tanner et al., 2014; Tanner & Van Hell, 2014).

2.4.1.1. <u>*Sentence stimuli-specific analyses*</u> The ERP responses of greatest interest from the sentence stimuli were the N400 response to sentences with semantic violations alone and sentences with double violations, and the P600 response to sentences with grammatical violations alone and sentences with double violations. To compute the effect magnitude of each of these responses, for each of the four sentence conditions, we calculated participants' mean ERP amplitudes in a central-posterior region of interest (electrodes C3, Cz, C4, P3, Pz, P4, O1, O2) where N400 and P600 effects are typically the largest (Tanner et al., 2014; Tanner & Van Hell, 2014). For sentences with semantic violations alone, semantic violation N400 effect magnitude was calculated as the mean activity in the well-formed minus semantic conditions between 300 and 500ms. For sentences with grammatical violations, grammatical violation P600 effect magnitude was calculated as the mean activity in the grammatical violation minus well-formed conditions between 500 and 900ms. For sentences with double violations, double violation N400 and P600 effect magnitudes were calculated in similar ways, comparing activity in the double violation and well-formed conditions. Thus, these calculations produced four ERP effect magnitudes: a) semantic violation N400, b) grammatical violation P600, c) double violation N400, and d) double violation P600.

These ERP effect magnitudes were compared to participants' standardized reading comprehension scores (WRMT-III Reading Comprehension scores, described earlier) in two ways. For all of these analyses, separate analyses were used for the deaf and hearing participants. First, we examined the simple correlations between the four ERP effect magnitudes and participants' standardized reading score. However, since many factors may lead to variation in a person's reading skill, we employed theoretically-driven multiple regression models to account for some of this variation and get a better picture of the actual relationship between these ERP measures and the participants' standardized reading skill.

As with the simple correlations, separate models were used for deaf and hearing participants. This was done for two reasons. First, language background was included as a predictor for deaf participants (see below), which was not a measure that existed for hearing participants. Second, since the relationships between predictors and outcomes was potentially different between the groups.

A schematic of the multiple regression models used for the sentence stimuli can be seen in Figure 4. The outcome measure of all multiple regression models was participants' standardized reading comprehension score. Predictors for each of the multiple regression models were of two kinds: a) background measures known to influence reading skill, and b) ERP measures of theoretical interest. Sequential predictor entry was used such that the background predictors



	Deaf	**Hearing**
	Predictors	
Background predictors (included in all models)	(1) Years of education (2) Standardized speechreading score (3) Language use while growing up (1-7 scale)	(1) Years of education (2) Standardized speechreading score
Model 1 ERP predictors	(4) Semantic violation N400 (5) Agreement violation P600	(3) Semantic violation N400 (4) Agreement violation P600
Model 2 ERP predictors	(4) Double violation N400 (5) Double violation P600	(3) Double violation N400 (4) Double violation P600

Outcome measure for all models: standardized reading comprehension score (WRMT-III Reading Comprehension Score)

Figure 4. Sentence stimuli multiple regression schematic.

were entered into the regression model first, and the ERP measures were entered second, allowing us to determine how much additional variance in standardized reading comprehension score the ERP measure accounted for. The background measures included as predictors for all participants, deaf and hearing, in each model were 1) years of education, and 2) standardized speechreading score. An additional background measure predictor was used in models with deaf participants: 3) their score on the 1-7 scale used to assess language use while growing up.

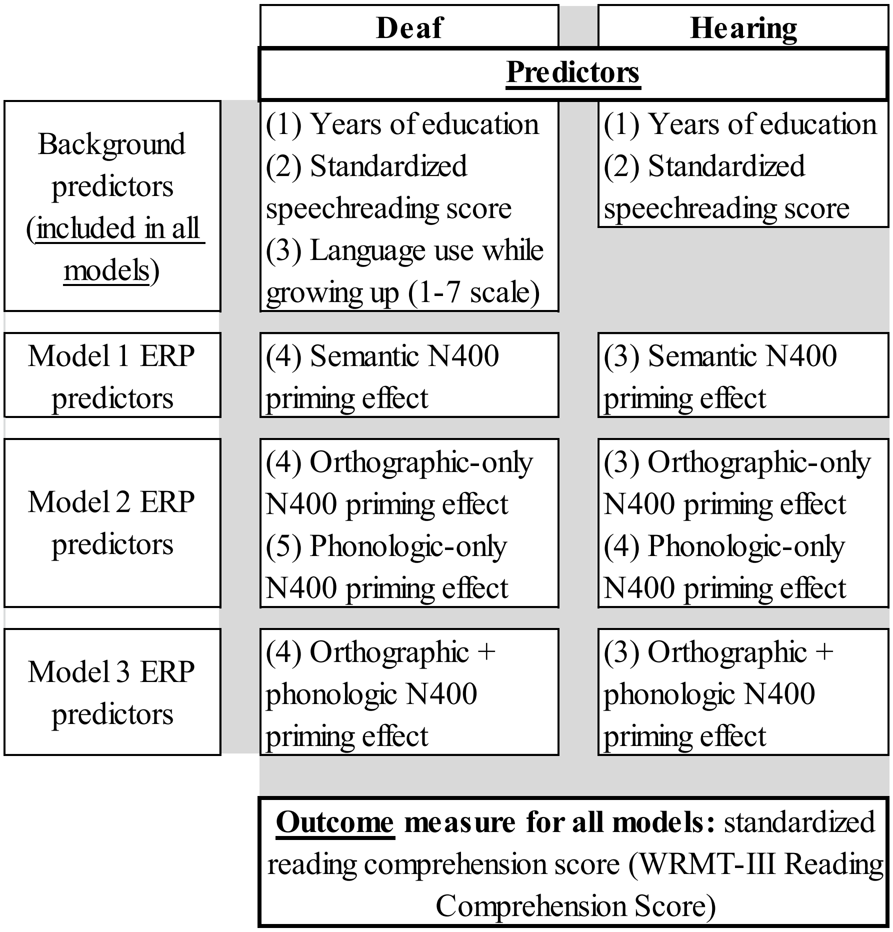
In regards to ERP measures as predictors, the four ERP effect magnitudes were split into two groups: 1) semantic violation N400 and grammatical violation P600, and 2) double violation N400 and P600. These two sets of ERP effect magnitudes were used as the final predictors in two separate multiple regression models (each with the same background measure predictors) for each group of participants (Figure 4). The design of these models was motivated by prior ERP language research that has shown that the N400 and P600 index different neural language processing streams (Kuperberg, 2007; Tanner, 2013).

These planned multiple regression analyses were used as the basis for power analysis used to determine the number of participants needed for this research study. With the assumption of a large effect size, a multiple regression model with five predictors and an alpha level $\alpha = 0.05$ requires 42 participants to obtain 80% power (Cohen, 1992). Thus, we recruited 42 deaf and 42 hearing participants for this study.

2.4.1.2. Word pair stimuli-specific analyses. The ERP response of interest from the word-pair stimuli was the magnitude of the reduction in N400 amplitude elicited by each target with a related prime, relative to the N400 elicited by targets with unrelated primes. Specifically, we were interested in the effect magnitude of the N400 reduction in response to: a) targets with semantically related primes, b) targets with orthographically related primes, c) targets with phonologically related primes, and d) targets whose primes were related in both orthography and phonology (recall Table 3 for examples of these stimuli). To compute the effect magnitude of each of these responses, for each of the four related targets and two unrelated targets, we calculated participants' mean ERP amplitudes between 300 and 500ms across the midline electrodes (electrodes Fz, Cz, and Pz). For semantically related word pairs, the semantic N400 priming effect magnitude was calculated as the mean activity in the semantically unrelated targets minus the mean activity in the semantically related targets. For word pairs related in only orthography, only phonology, or both orthography and phonology, N400 reduction effect magnitudes for each of these conditions was calculated as the mean activity in the orthographically/phonologically unrelated targets minus the mean activity in each of the three related targets. Thus, these calculations produced four ERP effect magnitudes: a) semantic N400 priming effect, b) orthographic-only N400 priming effect, c) phonologic-only N400 priming effect, and d) orthographic and phonologic N400 priming effect.

Because of the way the calculations were designed (unrelated minus related activity), a negative effect magnitude indicates that the N400 decreased in magnitude relative to the unrelated condition (i.e., a typical N400 reduction priming effect) and a positive effect magnitude indicates that the N400 increased in magnitude relative to the unrelated condition. Since the N400 is a negative-going ERP component, it is "largest" when it has the most negative amplitude. Discussing changes in N400 amplitude can be confusing because a "decrease in N400 magnitude" actually means that the N400 had a more positive voltage. Thus, we clarify to say that a "larger N400" means a more negative voltage, and a "smaller N400" means a more positive voltage.

Similar to the analyses used for the sentence stimuli, these ERP priming effect magnitudes were compared to participants' standardized reading comprehension scores in two ways. For all of these analyses, separate analyses were used for the deaf and hearing participants. First, we examined the simple correlations between the four ERP priming effect magnitudes and participants' standardized reading score.



	Deaf	Hearing
	Predictors	
Background predictors (included in all models)	(1) Years of education (2) Standardized speechreading score (3) Language use while growing up (1-7 scale)	(1) Years of education (2) Standardized speechreading score
Model 1 ERP predictors	(4) Semantic N400 priming effect	(3) Semantic N400 priming effect
Model 2 ERP predictors	(4) Orthographic-only N400 priming effect (5) Phonologic-only N400 priming effect	(3) Orthographic-only N400 priming effect (4) Phonologic-only N400 priming effect
Model 3 ERP predictors	(4) Orthographic + phonologic N400 priming effect	(3) Orthographic + phonologic N400 priming effect

Outcome measure for all models: standardized reading comprehension score (WRMT-III Reading Comprehension Score)

Figure 5. Word-pair stimuli multiple regression schematic.

Second, we employed theoretically-driven multiple regression models to account for other factors leading to variation in reading comprehension score. As with the simple correlations, separate models were used for deaf and hearing participants. A schematic of the multiple regression models used for the word-pair stimuli can be seen in Figure 5. Similar to the sentence stimuli multiple regression models, the outcome measure of all multiple regression models was participants' standardized reading comprehension score. The same background predictors and sequential entry order were used in the word-pair stimuli multiple regression models as were used in the sentence stimuli models.

In regards to ERP measures as predictors, the four ERP priming effect magnitudes were divided into three theoretically-driven groups: 1) semantic N400 priming effect, 2) orthographic-only N400 priming effect and phonologic-only N400 priming effect, and 3) orthographic and phonologic N400 priming effect. These three sets of ERP effect magnitudes were used as the final predictors in three separate multiple regression models (each with the same background measure predictors) for each group of participants (Figure 5). The design of these models allowed us to address a number of theoretically-driven questions. First, it allowed us to assess the relationship between semantic priming effects and reading comprehension. Second, it allowed us to examine the interplay between how orthographic and phonologic priming responses may be related to reading comprehension. Finally, it allowed us to look at the relationship between combined orthographic and phonologic priming alone. The same power analysis used for the sentence stimuli multiple regression models was also relevant here; as no more than five predictors were used in any of the models.

Chaper 3. Results

3.1. Sentence data

3.1.1. End-of-sentence acceptability judgment task. To analyze responses to the end-of-sentence acceptability judgment task, *d'* scores were calculated for each participant for each of the three types of sentence violations: grammatical violations, semantic violations, and double violations. An overall *d'* score was also calculated. Average *d*' scores for all participants and conditions can be seen in Figure 6. Four t-tests (corrected alpha $\alpha = 0.0125$) were used to compare deaf and hearing participants' performance across the three sentence conditions and overall. For all three types of sentence violations and overall, hearing participants were better at discriminating sentence violations than deaf participants ($t(82)$ ranged from 4.070 to 5.111, all p's < 0.001). To compare performance on each type of sentence condition within each group, one-way ANOVAs with Tukey post-hoc tests were used to analyze each group's *d*' scores

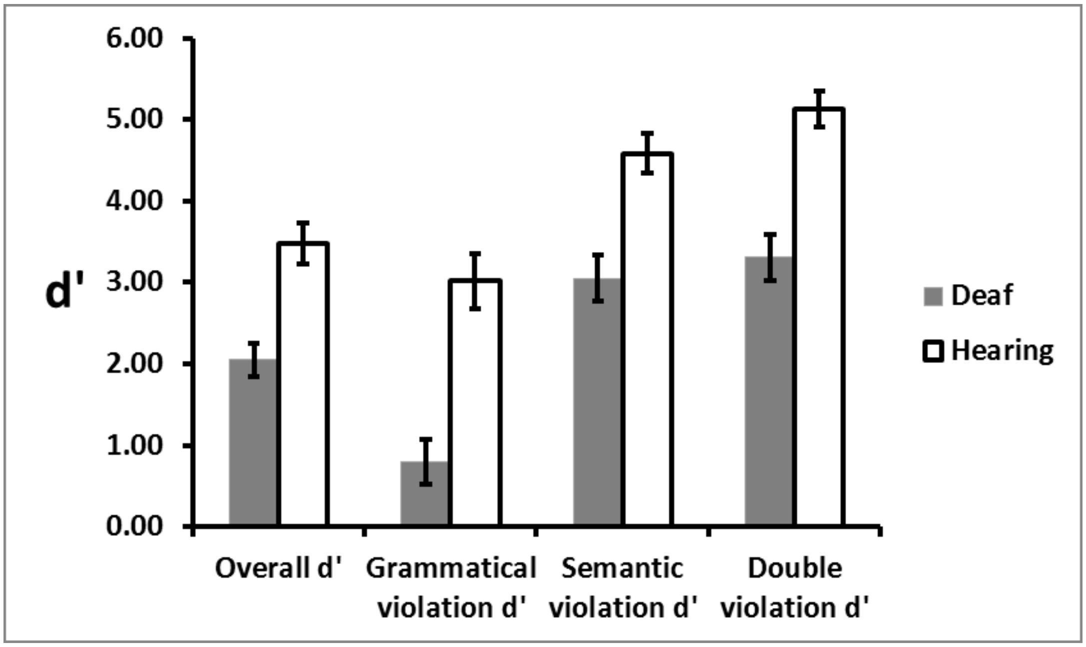


Figure 6. End-of-sentence acceptability judgment *d*' scores for deaf and hearing participants. Overall *d*' refers to performance across all types of sentence violations. Error bars represent one standard error.

separately. For both deaf and hearing participants, there was a significant difference in *d*' score between the different sentence conditions (deaf: $F(2,123) = 23.982, p < 0.001$; hearing: $F(2, 123) = 16.202, p < 0.001$). Tukey post-hoc tests determined that in both groups, *d*' scores were lower for sentences with grammatical violations than for sentences with semantic or double violations (all *p*'s < 0.001). For both groups, there was no significant difference in the *d*' scores between sentences with semantic and double violations. In sum, the hearing participants were better at discriminating sentence acceptability than deaf participants across all types of sentence violations, but both groups showed similar trends of *d*' differences between different types of sentence violations.

For both deaf and hearing participants, all *d*' scores were significantly positively correlated with participants' reading comprehension score; as any *d*' score increased, so did reading comprehension scores. R^2 values for these correlations can be seen in Table 4. Of particular note is that for deaf participants, while the grammatical violation *d*', semantic violation *d*', and double violation *d*' all explain a significant amount of variance in reading comprehension score, the semantic and double violation *d*' scores each explain over four times as much variance in reading comprehension than the grammatical violation *d*' score does.

Table 4. Coefficients of determination (r^2) for the relationship between end-of-sentence acceptability judgment d' scores and standardized reading comprehension scores.

	Standardized reading comprehension score	
	Deaf	**Hearing**
Overall *d*'	0.50 ***	0.22 **
Grammatical violation *d*'	0.12 *	0.32 ***
Semantic violation *d*'	0.52 ***	0.16 **
Double violation *d*'	0.56 ***	0.29 ***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: all correlation coefficients (r) are positive.

3.1.2. Grand mean results. For both the deaf and hearing groups, grand mean ERP results for semantic violations can be seen in Figure 7, for grammatical violations in Figure 8, and for double violations in Figure 9. Data from the 300-500ms (N400) and 500-900ms (P600) time windows were analyzed separately.

3.1.2.1. *N400 (300-500ms) time window.* Visual inspection of the grand mean ERP waveforms showed that relative to well-formed sentences, sentences with a semantic violation (Figure 7) or a double violation (Figure 9) elicited a widely-distributed negativity with a posterior maximum between approximately 300 and 500ms (N400) in both deaf and hearing participants. There did not appear to be differences in this time window when comparing well-formed sentences and sentences with a grammatical violation.

Statistical analyses confirmed these observations. Reflecting the widespread N400 in response to semantic and double violations, there was a main effect of semantic correctness (midline: $F(1,82) = 58.026$, $p < 0.001$; medial: $F(1,82) = 55.105$, $p < 0.001$; lateral: $F(1,82) = 21.637$, $p < 0.001$). This effect was strongest over posterior electrodes, (semantic correctness x electrode interaction: midline: $F(2,164) = 25.371$, $p < 0.001$; medial: $F(4,328) = 18.338$, $p < 0.001$; lateral: $F(2,164) = 25.269$, $p < 0.001$).

There was also an interaction between semantic correctness and grammaticality at midline and lateral electrode sites (semantic correctness x grammaticality interaction: midline: $F(1,82) = 4.239$, $p = 0.043$; lateral: $F(1,82) = 4.039$, $p = 0.048$). Follow-up contrasts (corrected alpha $\alpha = 0.025$ for each scalp region) indicated that this interaction was driven by two factors. First, the interaction was driven by the lack of N400 change elicited by grammatical violations alone, as compared to the much larger N400 elicited by double violations (midline: $F(1,83) = 45.666$, $p < 0.001$; lateral: $F(1,83) = 28.134$, $p < 0.001$).

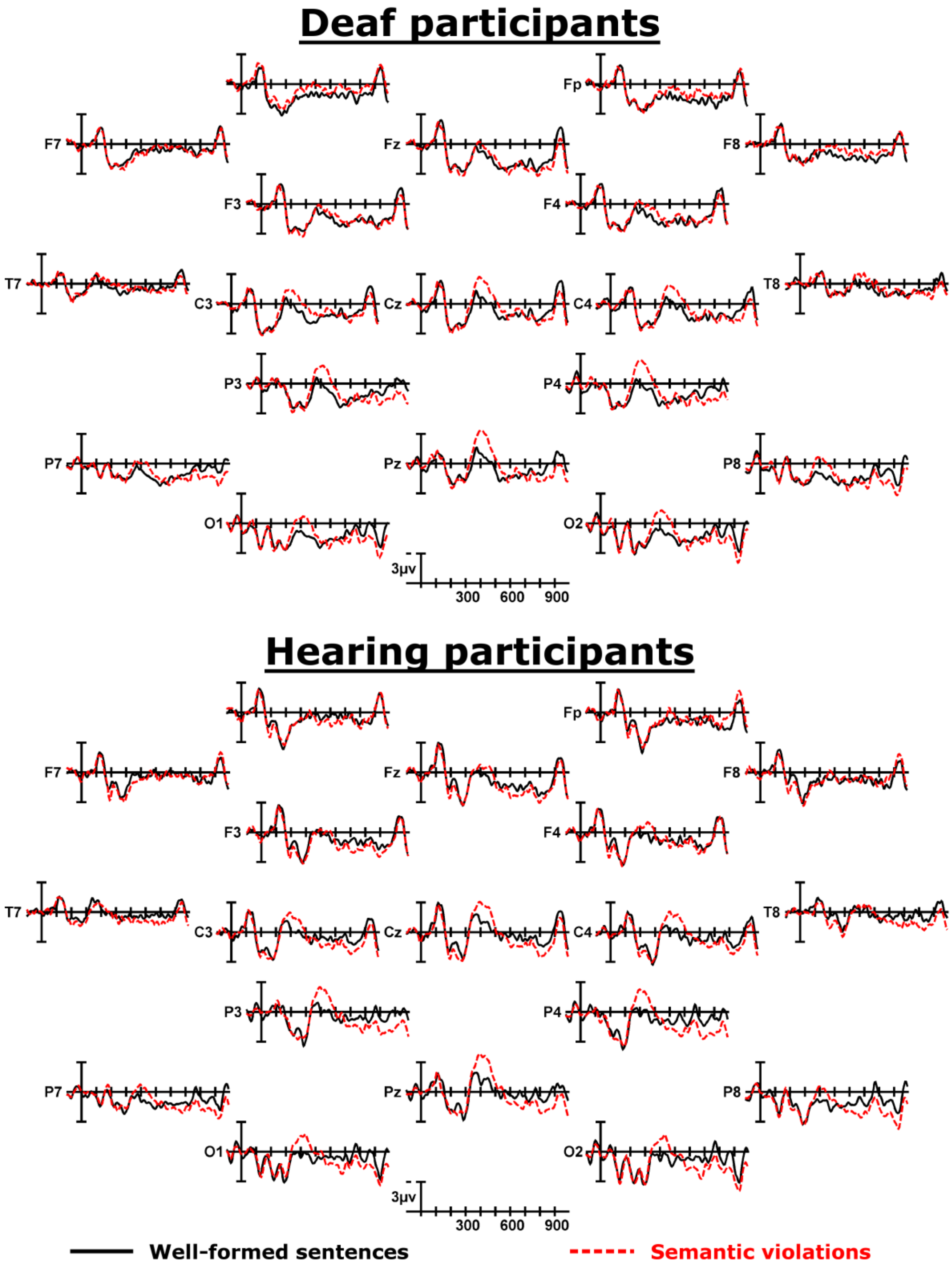


Figure 7. Grand mean ERP waveforms for sentences with semantic violations (red dashed line) and well-formed sentences (solid black line) for deaf (top) and hearing (bottom) participants. Onset of the critical word in the sentence is indicated by the vertical bar. Calibration bar shows 3µVof activity; each tick mark represents 100ms of time. Negative voltage is plotted up.

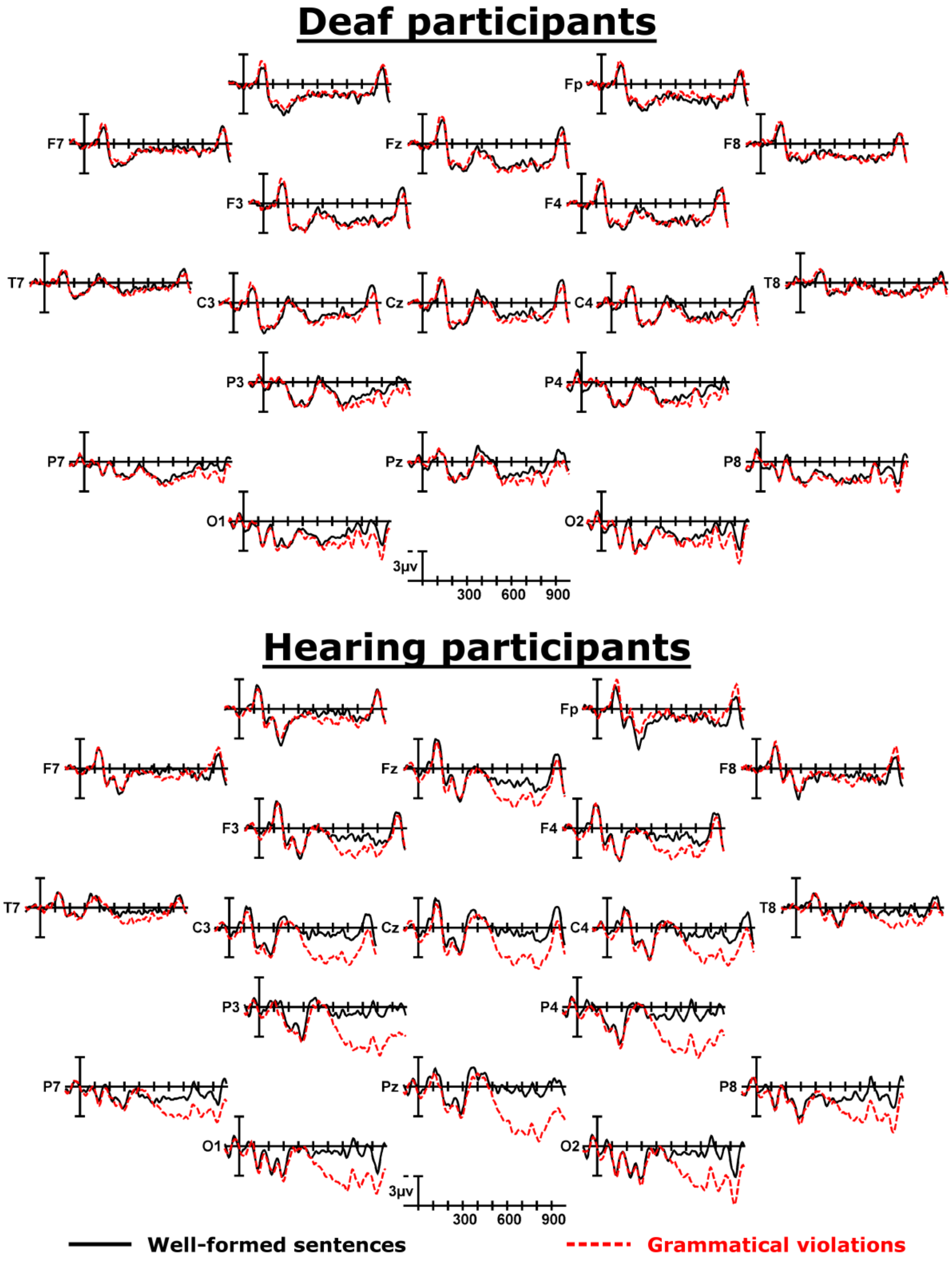


Figure 8. Grand mean ERP waveforms for sentences with grammatical violations (red dashed line) and well-formed sentences (solid black line) for deaf (top) and hearing (bottom) participants. Onset of the critical word in the sentence is indicated by the vertical bar. Calibration bar shows 3µVof activity; each tick mark represents 100ms of time. Negative voltage is plotted up.

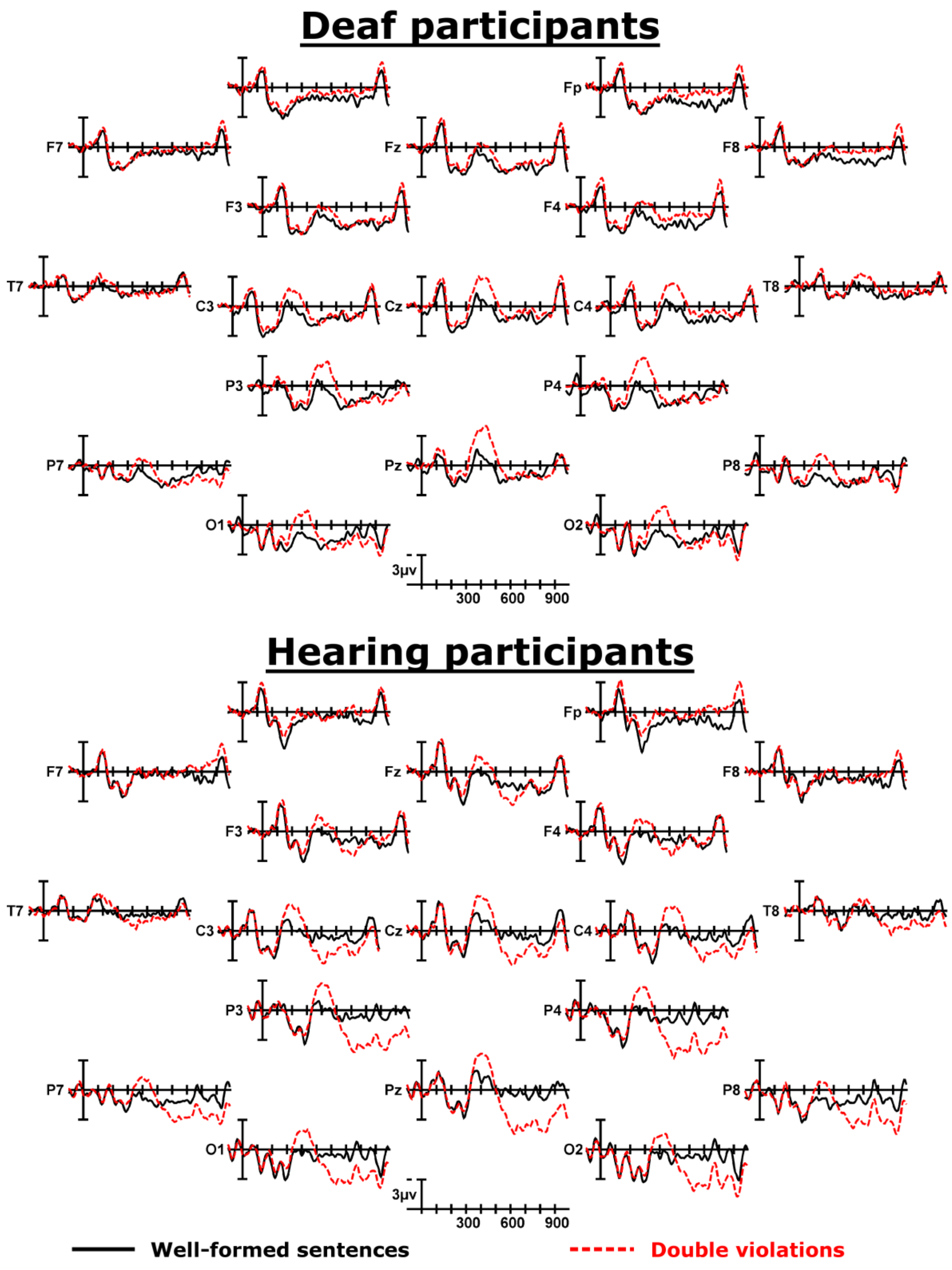


Figure 9. Grand mean ERP waveforms for sentences with double (semantic and grammatical) violations (red dashed line) and well-formed sentences (solid black line) for deaf (top) and hearing (bottom) participants. Onset of the critical word in the sentence is indicated by the vertical bar. Calibration bar shows 3μVof activity; each tick mark represents 100ms of time. Negative voltage is plotted up.

Second, at lateral electrode sites the N400 response to double violations was significantly larger than the N400 response to semantic violations alone (midline: $F(1,83) = 3.588$, $p = 0.062$; lateral: $F(1,83) = 7.132$, $p = 0.009$)

Deaf and hearing participants did not differ significantly in any of the previously reported analyses; there were no group differences in N400 responses to semantic and double violations in sentences.

Small hemispheric differences in the N400 response at lateral electrodes were observed (semantic correctness x hemisphere: lateral: $F(1,82) = 4.384$, $p = 0.039$), which was driven by an interaction between the groups (semantic correctness x hemisphere x group interaction: lateral: $F(1,82) = 4.041$, $p = 0.048$). Follow-up ANOVAs with the deaf and hearing participants in separate models (corrected alpha $\alpha = 0.025$) indicated that this interaction was due to deaf participants having larger N400s in the right hemisphere than in the left (deaf participants: semantic correctness x hemisphere interaction: lateral: $F(1,41) = 7.114$, $p = 0.011$) while hearing participants did not show any significant hemispheric differences.

Neither group of participants showed any evidence of a LAN in response to grammatical violations (Figure 8).

In summary, in the 300-500ms time window, both deaf and hearing participants showed large N400 responses to sentences with semantic or double violations, and no significant response to sentences with grammatical violations.

3.1.2.2. *P600 (500-900ms) time window.* Visual inspection of the grand mean ERP waveforms showed marked differences between hearing and deaf participants in the 500-900ms time window. In hearing participants, relative to well-formed sentences, sentences with a grammatical violation (Figure 8) or a double violation (Figure 9) elicited a large, widely-distributed positivity beginning around 500ms with a posterior maximum (a P600 effect). Deaf participants showed only a small evidence of this positivity in response to sentences with grammatical or double violations (Figure 8, Figure 9). Both groups of participants showed a small positivity in response to sentences with semantic violations (Figure 7).

Statistical analyses confirmed these observations. Across all participants, there was a main effect of grammaticality (midline: $F(1,82) = 10.181$, $p = 0.002$; medial: $F(1,82) = 10.414$, $p = 0.002$; lateral: $F(1,82) = 3.779$, $p = 0.055$). The interpretation of the main effect of grammaticality, however, is conditional upon the fact that there was an interaction between group (deaf vs. hearing) and grammaticality (grammaticality x group interaction: midline: $F(1,82) = 10.413$, $p = 0.002$; medial: $F(1,82) = 8.440$, $p = 0.005$; lateral: $F(1,82) = 4.108$, $p = 0.046$). Follow-up ANOVAs with the deaf and hearing participants in separate models (corrected alpha $\alpha = 0.025$ for each scalp region) showed that for hearing participants, relative to grammatical sentences, sentences containing a grammatical or double violation elicited a robust P600 (hearing participants: grammaticality main effect: midline: $F(1,41) = 14.637$, $p < 0.001$; medial: $F(1,41) = 13.194$, $p = 0.001$; lateral: $F(1,41) = 6.156$, $p = 0.017$), whereas there was no main effect of grammaticality for deaf participants alone.

Across all participants, the main effect of grammaticality was largest over posterior electrodes (grammaticality x electrode interaction: midline: $F(2,164) = 27.520$, $p < 0.001$; medial: $F(4,328) = 21.806$, $p < 0.001$; lateral: $F(2,164) = 18.514$, $p < 0.001$), but again, interpreting this effect is dependent on the fact that there were differences between the two groups in this interaction (grammaticality x electrode x group interaction: midline: $F(2,164) = 11.380$, $p < 0.001$; medial: $F(4,328) = 8.087$, $p = 0.001$; lateral: $F(2,164) = 7.200$, $p = 0.005$). Follow-up ANOVAs with the deaf and hearing participants in separate models (corrected alpha α

= 0.025 for each scalp region) indicated that in hearing participants, the P600 elicited by sentences with grammatical and double violations was largest at posterior electrodes (hearing participants: grammaticality x electrode interaction: midline: $F(2,82) = 32.445$, $p < 0.001$; medial: $F(4,164) = 22.892$, $p < 0.001$; lateral: $F(2,82) = 20.868$, $p < 0.001$), whereas this interaction was not present in the deaf participants alone.

The deaf and hearing participants showed different hemispheric effects of grammaticality (grammaticality x hemisphere x group interaction: medial: $F(1,82) = 4.423$, $p = 0.039$; lateral: $F(1,82) = 5.110$, $p = 0.026$). Follow-up ANOVAs with the deaf and hearing participants in separate models (corrected alpha $\alpha = 0.025$ for each scalp region) indicated that in deaf participants, there were hemispheric differences in the effect of grammaticality at medial electrodes (deaf participants: grammaticality x hemisphere interaction: medial: $F(1,41) = 6.282$, $p = 0.016$). Hemispheric interactions at lateral electrode sites did not reach significance, and there were no significant hemispheric interactions with grammaticality in the hearing participants alone.

While these was no main effect of semantics in the 500-900ms time window, all participants had a significant P600-like response at posterior medial and lateral electrode sites to sentences containing a semantic violation (whether it be a semantic violation alone or a double violation) (semantic correctness x electrode interaction: medial: $F(4,328) = 5.735$, $p = 0.006$; lateral: $F(2,164) = 14.157$, $p < 0.001$)

Finally, there was also an interaction between semantic correctness and grammaticality in all participants (semantic correctness x grammaticality interaction: midline: $F(1,82) = 12.629$, $p = 0.001$; medial: $F(1,82) = 10.413$, $p = 0.002$; lateral: $F(1,82) = 9.164$, $p = 0.003$). Follow-up contrasts comparing semantic and grammatical violation sentences alone with the double violation sentences (corrected alpha $\alpha = 0.025$ for each scalp region) indicated that this interaction was driven by two factors. First, the P600 elicited by sentences with grammatical violations alone was significantly larger than the P600 elicited by sentences with double violations (midline: $F(1,83) = 9.597$, $p = 0.003$; medial: $F(1,83) = 9.484$, $p = 0.003$; lateral: $F(1,83) = 5.268$, $p = 0.024$). This is likely due to the N400 response also elicited by double violations, the presence of which pulls the P600 "up" – more negative/less positive. While deaf participants did not display widely-distributed P600s in response to grammatical or double violations, this trend is still present in their responses when comparing the relative voltages of the 500-900ms time window between the response to grammatical violations alone versus the response to double violations. Second, at posterior electrodes, the P600 elicited by semantic violations alone was smaller than the P600 elicited by double violations, while this trend reversed at anterior electrodes (electrode interaction: midline: $F(2,166) = 20.109$, $p < 0.001$; medial: $F(4,332) = 9.582$, $p < 0.001$; lateral: $F(2,166) = 9.508$, $p = 0.001$).

In summary, in the 500-900ms time window, hearing participants displayed large P600s in response to sentences with grammatical or double violations, while deaf participants did not show evidence of significant P600s as a group. Both groups of participants displayed P600-like responses to semantic violations in sentences.

3.1.3. Individual differences analyses. As described in the methods section, effect magnitudes of the following ERP responses were calculated for use in sentence-level individual differences analyses: a) semantic violation N400, b) grammatical violation P600, c) double violation N400, and d) double violation P600.

Simple correlations between the four ERP effect magnitudes and participants' standardized reading comprehension scores can be seen in Table 5. For deaf participants, as the magnitude of the N400 elicited by sentences with double violations increased, reading comprehension score also increased. No other ERP effect magnitude was significantly correlated with reading comprehension score for the deaf participants. For hearing participants, however, a notably different set of relationship was found. In the hearing participants, as the magnitude of the P600 elicited by sentences with either a grammatical violation alone or a double violation increased, reading comprehension score increased. The relationship between the magnitude of the N400 elicited by semantic violations alone and reading comprehension score approached significance for hearing participants, with a larger semantic N400 effect magnitude being associated with higher reading comprehension score.

Table 5. Correlation coefficients (*r*) for the relationship between sentence stimuli ERP effect magnitudes and standardized reading comprehension score.

	Standardized reading comprehension score	
	Deaf	**Hearing**
Semantic violation N400	0.161	0.291^
Grammatical violation P600	0.193	0.339*
Double violation N400	0.310*	0.098
Double violation P600	0.151	0.323*

$^{*} p < 0.05$; $^{\wedge} p < 0.1$

As was described in the methods, a number of factors may lead to variation in a person's reading skill. To better understand the relationship between the ERP effect magnitudes and participants' standardized reading comprehension scores, we employed theoretically-driven multiple regression models to account for some of this variation, previously described in Figure 4. As the schematic shows, the first set of multiple regression models included as ERP predictors the effect magnitudes of the semantic violation N400 and grammatical violation P600. The results of these models, one model for deaf and one for hearing participants, can be seen in Table 6.

Table 6. Sentence stimuli multiple regression models using semantic violation N400 and grammatical violation P600 as the ERP predictors (Model 1).

	Deaf Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.64*	*22.12* *(3,38)****	*.64*	*.61*	*22.12* *(3,38)****					
Years of Education						3.73	(1.07)	0.41	3.50 **	0.001
NTID Speechreading						0.47	(0.11)	0.53	4.10 ***	<0.001
Grow Up Language						0.81	(0.99)	0.09	0.82	0.419
Block 2	*.05*	*2.98* *(2,36)^*	*.69*	*.64*	*15.85* *(5,36)****					
Years of Education						4.11	(1.03)	0.45	3.98 ***	<0.001
NTID Speechreading						0.44	(0.11)	0.51	4.03 ***	<0.001
Grow Up Language						0.22	(0.98)	0.03	0.23	0.821
Semantic Viol. N400						2.38	(0.98)	0.23	2.42 *	0.021
Grammatical Viol. P600						-0.29	(1.03)	-0.03	-0.28	0.778

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* $p < .05$, ** $p < .01$, *** $p < .001$. ^$p < 0.1$

	Hearing Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.17*	*3.99* *(2,39)**	*.17*	*.13*	*3.99* *(2,39)**					
Years of Education						1.83	(0.72)	0.37	2.56 *	0.014
NTID Speechreading						0.14	(0.13)	0.16	1.09	0.284
Block 2	*.17*	*4.79* *(2,37)**	*.34*	*.27*	*4.78* *(4,37)***					
Years of Education						1.47	(0.67)	0.30	2.18 *	0.036
NTID Speechreading						0.16	(0.12)	0.18	1.29	0.205
Semantic Viol. N400						1.72	(0.86)	0.28	2.00 ^	0.053
Grammatical Viol. P600						1.40	(0.56)	0.34	2.53 *	0.016

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* $p < .05$, ** $p < .01$, *** $p < .001$. ^$p < 0.1$

Our focus is less on the overall model fit and more on the predictive value of each individual predictor, as well as how much additional variance in reading comprehension score the ERP measures account for. For both deaf and hearing participants, an increase in the number of years of education completed predicted a higher reading comprehension score. For only the deaf participants, a higher score on the NTID speechreading test predicted a higher reading comprehension score. Notably, our measure of language use while growing up did not significantly predict reading comprehension score for deaf participants.

In terms of the ERP predictors, for deaf participants, an increase in the effect magnitude of the N400 elicited by semantic violations alone predicted a higher reading comprehension score. The effect magnitude of the P600 elicited by grammatical violations alone was not a predictor of reading comprehension score for deaf participants. Conversely, in hearing participants, an increase in the effect magnitude of the P600 elicited by grammatical violations did significantly predict higher reading comprehension scores. The effect magnitude of the N400 elicited by semantic violations approached significance in predicting hearing participants' reading comprehension scores, with a semantic N400 effect magnitude predicting higher reading comprehension score.

The second set of regression models included as ERP predictors the effect magnitudes of the double violation N400 and P600 (recall the schematic in Figure 4). The results of these models, one model for deaf and one for hearing participants, can be seen in Table 7. As was seen in the first set of models, for both deaf and hearing participants, an increase in the number of years of education completed predicted a higher reading comprehension score. For only the deaf participants, a higher score on the NTID speechreading test predicted a higher reading comprehension score.

Table 7. Sentence stimuli multiple regression models using double violation N400 and P600 as the ERP predictors (Model 2).

	Deaf Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.64*	*22.12 (3,38)****	*.64*	*.61*	*22.12 (3,38)****					
Years of Education						3.73	(1.07)	0.41	3.50 **	0.001
NTID Speechreading						0.47	(0.11)	0.53	4.10 ***	<0.001
Grow Up Language						0.81	(0.99)	0.09	0.82	0.419
Block 2	*.14*	*10.76 (2,36)****	*.77*	*.74*	*24.39 (5,36)****					
Years of Education						4.17	(0.91)	0.46	4.57 ***	<0.001
NTID Speechreading						0.38	(0.10)	0.43	3.98 ***	<0.001
Grow Up Language						-0.01	(0.82)	0.00	-0.01	0.990
Double Viol. N400						3.69	(0.89)	0.45	4.16 ***	<0.001
Double Viol. P600						1.10	(0.99)	0.13	1.11	0.273

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* *p* < .05, ** *p* < .01, *** *p* < .001. ^ *p* < 0.1

	Hearing Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.17*	*3.99 (2,39)**	*.17*	*.13*	*3.99 (2,39)**					
Years of Education						1.83	(0.72)	0.37	2.56 *	0.014
NTID Speechreading						0.14	(0.13)	0.16	1.09	0.284
Block 2	*.14*	*3.74 (2,37)**	*.31*	*.24*	*4.15 (4,37)***					
Years of Education						1.61	(0.68)	0.33	2.39 *	0.022
NTID Speechreading						0.15	(0.13)	0.17	1.18	0.247
Double Viol. N400						1.83	(0.99)	0.29	1.84 ^	0.073
Double Viol. P600						1.51	(0.58)	0.39	2.59 *	0.014

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* *p* < .05, ** *p* < .01, *** *p* < .001. ^ *p* < 0.1

Similar relationships between ERP predictors and reading comprehension as were seen in the first model were found here. For deaf participants, an increase in the effect magnitude of the N400 elicited by double violations predicted a higher reading comprehension score. The effect magnitude of the P600 elicited by double violations was not a predictor of reading comprehension score for deaf participants. Conversely, in hearing participants, an increase in the effect magnitude of the P600 elicited by double violations did significantly predict higher reading comprehension scores. The effect magnitude of the N400 elicited by double violations approached significance in predicting hearing participants' reading comprehension scores, again with a larger semantic N400 effect magnitude predicting higher reading comprehension score.

ERP effect magnitudes were also compared to participants *d*’ scores from each type of sentence violation (Table 8). For deaf participants, none of the ERP effect magnitudes were significantly correlated with their corresponding *d*’ score. For hearing participants, the effect magnitude of the grammatical violation P600 was significantly correlated with grammatical violation *d*’ score; as the grammatical violation P600 effect magnitude increased, the grammatical violation *d*’ score also increased. A similar but approaching-significance relationship was seen in hearing participants for the semantic violation N400 effect magnitude and semantic violation *d*’ score.

Table 8. Correlation coefficients (*r*) for the relationship between sentence stimuli ERP effect magnitudes and d' scores.

Correlation	Deaf	Hearing
Semantic violation N400 vs. semantic violation *d*'	0.078	0.297^
Grammatical violation P600 vs. grammatical violation *d*'	0.143	0.389*
Double violation N400 vs. double violation *d*'	0.205	0.234
Double violation P600 vs. double violation *d*'	0.221	0.033

** $p < 0.05$; ^ $p < 0.1$*

Overall, these multiple regression models show that for the deaf participants, the magnitude of the N400 response to semantic violations alone or double violations was the best ERP predictor of reading comprehension score, with larger N400s predicting better reading scores. For the hearing readers, the magnitude of the P600 response to grammatical violations alone or double violations was the best ERP predictor of reading comprehension score, with larger P600s predicting better reading. Hearing participants also appear to have a weaker but approaching-significance relationship between N400 effect magnitudes and reading comprehension score, with larger N400s predicting better reading comprehension scores.

3.2. Word pair data

3.2.1. Lexical decision task. While participants were asked to perform a lexical decision task while reading the word pairs, this task was not directly related to the ERP responses of interest. Thus, participants' performance on the lexical decision task is not reported here.

3.2.2. Grand mean results. All figures with ERP waveforms show the response to the target (second) word in a word pair.

3.2.2.1. Semantically related word pairs. For both deaf and hearing participants, grand mean ERP results for semantically related word pairs can be seen in Figure 10. Visual inspection of the grand mean ERP waveforms show that relative to target words with an unrelated prime, target words with a semantically related prime elicited a widely-distributed reduction in N400 amplitude, in both deaf and hearing participants.

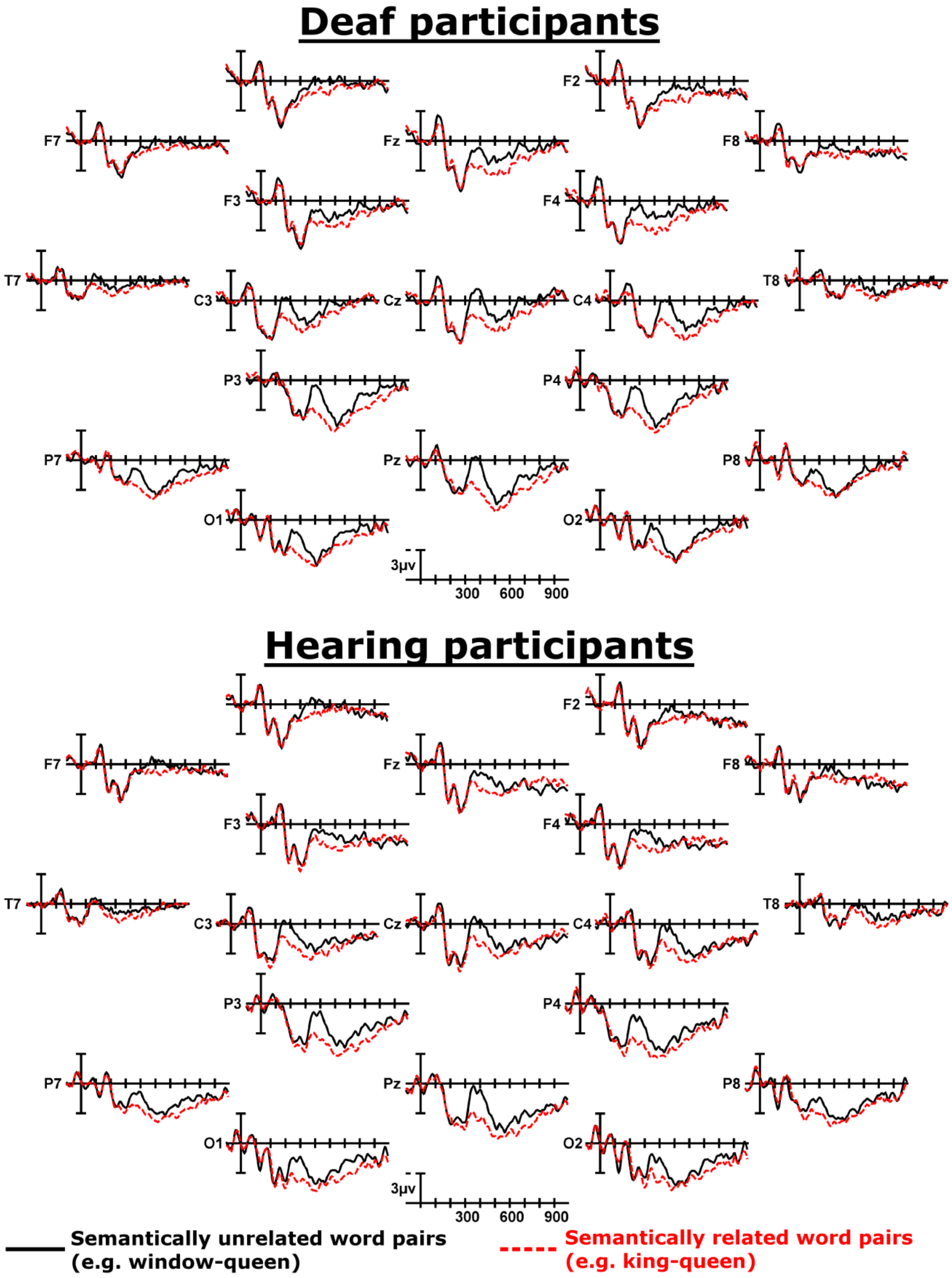


Figure 10. Grand mean ERP waveforms for semantically related word pairs (red dashed line) and unrelated word pairs (solid black line) for deaf (top) and hearing (bottom) participants. Onset of the target word in the pair is indicated by the vertical bar. Calibration bar shows 3µVof activity; each tick mark represents 100ms of time. Negative voltage is plotted up.

Statistical analyses confirmed these observations. There was a main effect of semantic relatedness (midline: $F(1,82) = 56.682$, $p < 0.001$; medial: $F(1,82) = 54.357$, $p < 0.001$; lateral: $F(1,82) = 25.741$, $p < 0.001$), reflecting the widely-distributed reduction in N400 amplitude in response to semantically related word pairs. The N400 reduction was strongest over posterior electrodes (semantic relatedness x electrode interaction: midline: $F(2,164) = 11.537$, $p < 0.001$; medial: $F(4,328) = 10.815$, $p < 0.001$; lateral: $F(2,164) = 12.509$, $p < 0.001$). There were no differences between the deaf and hearing participants in the responses to semantically related word pairs; both groups showed similar reductions in N400 amplitude in response to semantically related word pairs.

3.2.2.2. Orthographically and phonologically related word pairs. For both deaf and hearing participants, grand mean ERP results for orthographically-only related word pairs can be seen in in Figure 11, for phonologically-only related word pairs in Figure 12, and for word pairs related in both orthography and phonology in Figure 13. Visual inspection of the grand mean ERP waveforms showed that relative to target words with an unrelated prime, target words with a prime related in orthography alone or in both orthography and phonology elicited a widely-distributed reduction in N400 amplitude in both deaf and hearing participants. Words related in both orthography and phonology elicited slightly more of a reduction in N400 amplitude than words related only in orthography. Words related in phonology elicited only a small reduction in N400 amplitude in hearing participants, and no N400 reduction in deaf participants.

Statistical analyses confirmed these observations. There was a main effect of orthographic relatedness (midline: $F(1,82) = 66.071$, $p < 0.001$; medial: $F(1,82) = 68.631$, $p < 0.001$; lateral: $F(1,82) = 44.844$, $p < 0.001$),

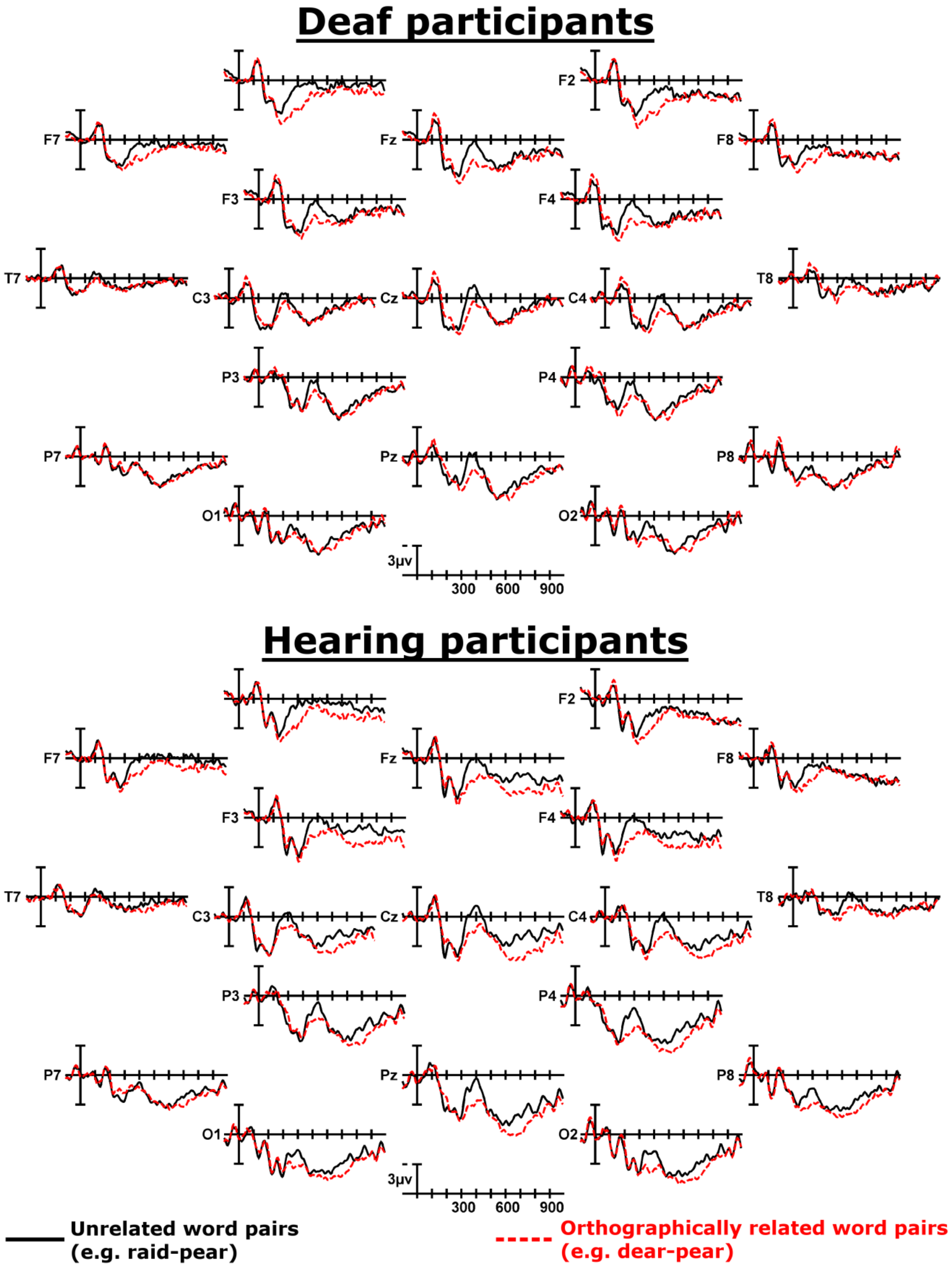


Figure 11. Grand mean ERP waveforms for orthographically related word pairs (red dashed line) and unrelated word pairs (solid black line) for deaf (top) and hearing (bottom) participants. Onset of the target word in the pair is indicated by the vertical bar. Calibration bar shows 3µVof activity; each tick mark represents 100ms of time. Negative voltage is plotted up.

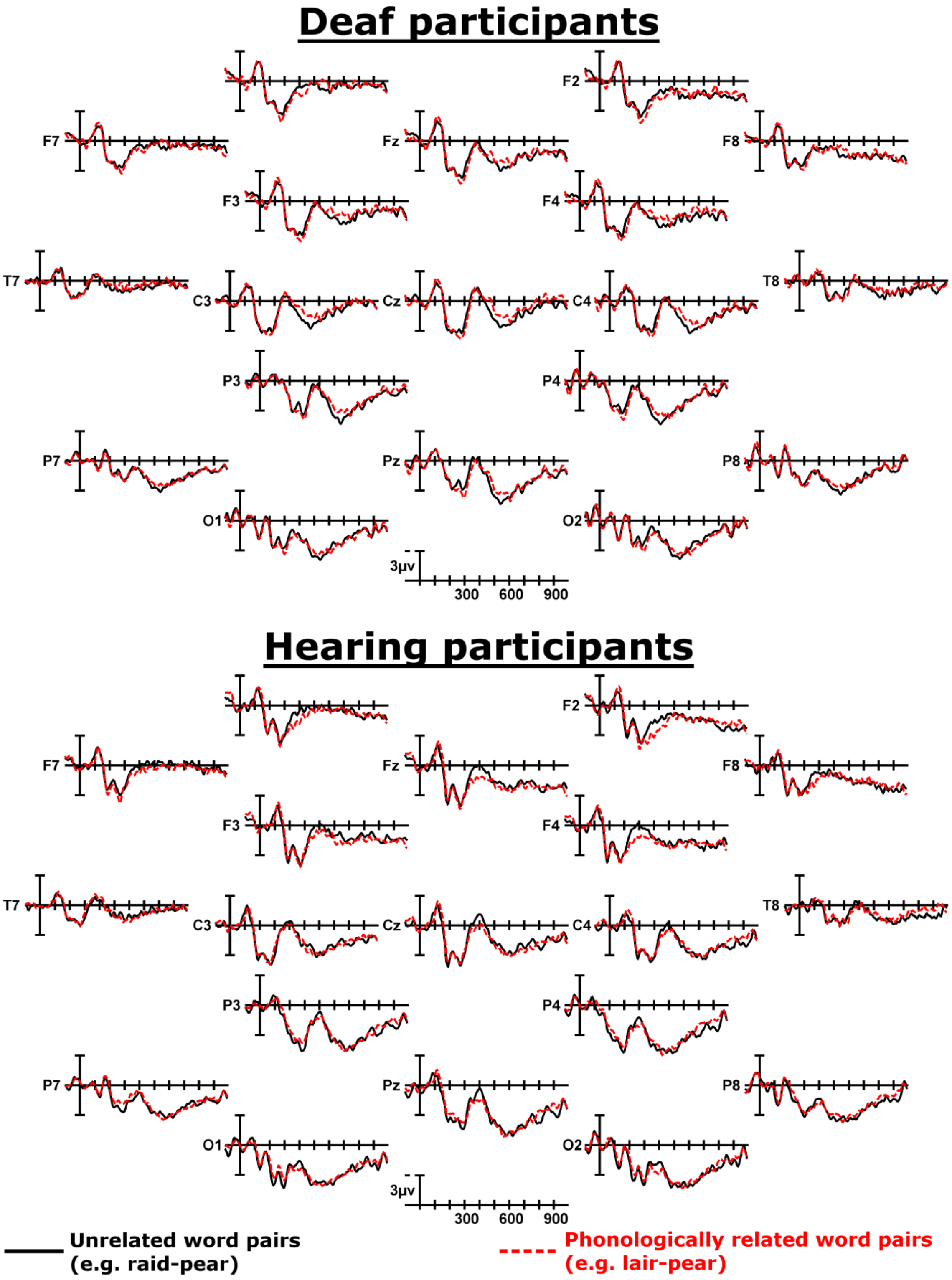
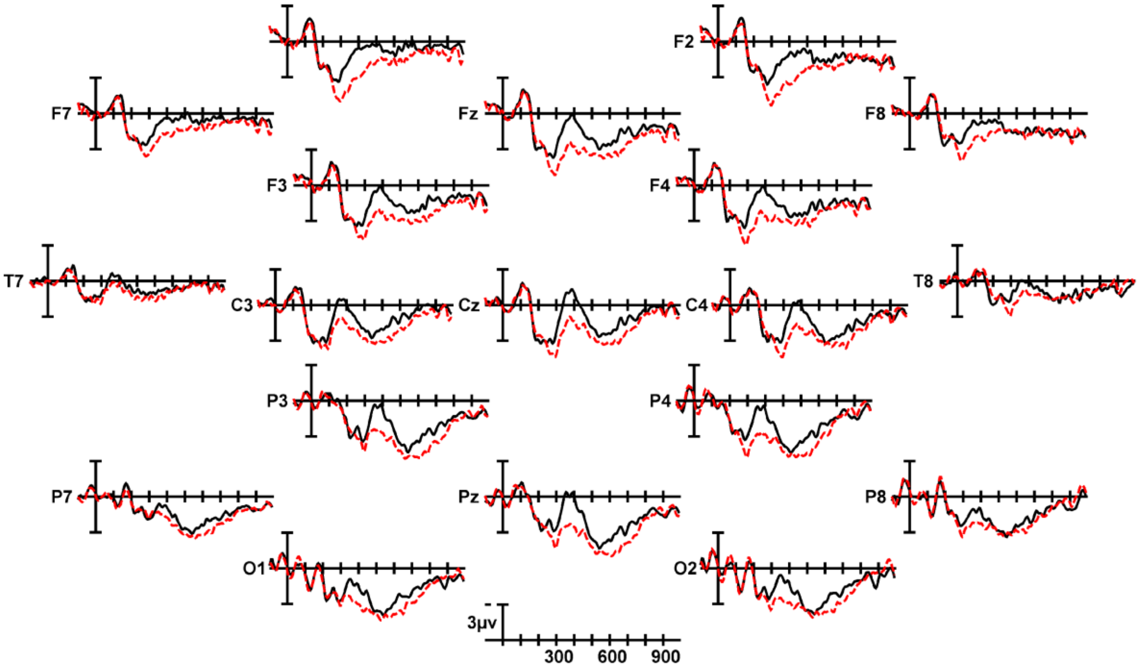


Figure 12. Grand mean ERP waveforms for phonologically related word pairs (red dashed line) and unrelated word pairs (solid black line) for deaf (top) and hearing (bottom) participants. Onset of the target word in the pair is indicated by the vertical bar. Calibration bar shows 3µVof activity; each tick mark represents 100ms of time. Negative voltage is plotted up.

Deaf participants



Hearing participants

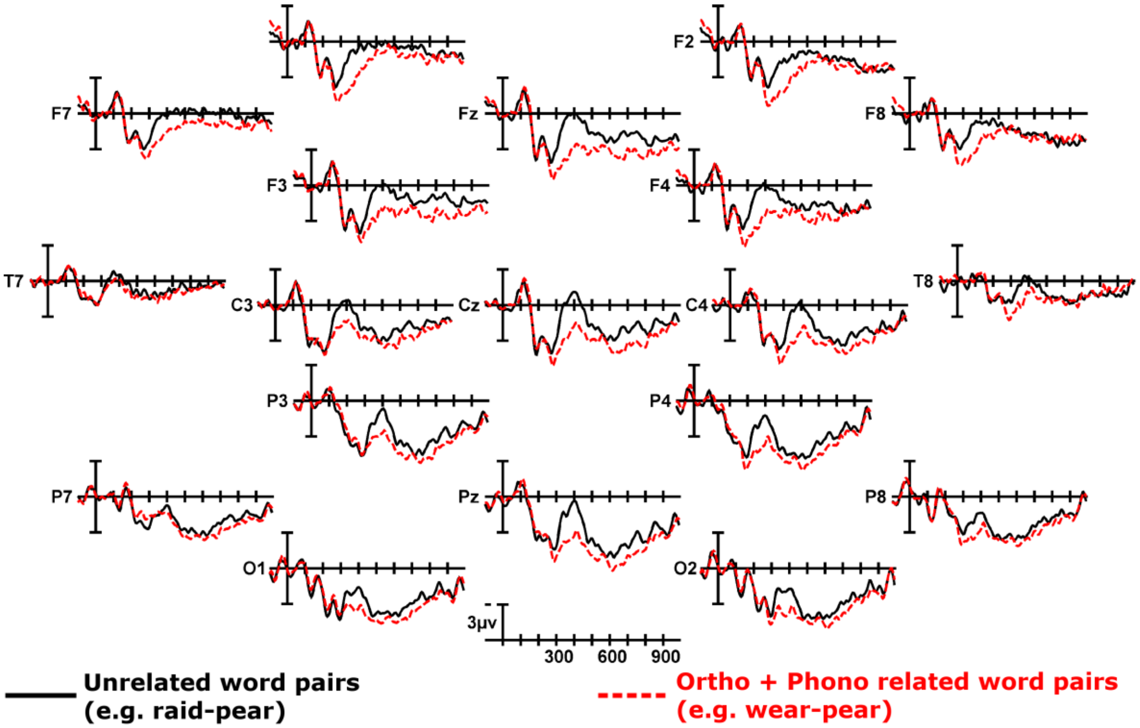


Figure 13. Grand mean ERP waveforms for word pairs related in both orthography and phonology (red dashed line) and unrelated word pairs (solid black line) and for deaf (top) and hearing (bottom) participants. Onset of the target word in the pair is indicated by the vertical bar. Calibration bar shows 3µVof activity; each tick mark represents 100ms of time. Negative voltage is plotted up.

reflecting the widespread reduction in N400 amplitude elicited by word pairs related in either only orthography or in both orthography and phonology. At medial and lateral electrode sites, this N400 reduction was strongest in frontal electrodes (orthographic relatedness x electrode interaction: medial: $F(4,328) = 6.628$, $p = 0.002$; lateral: $F(2,164) = 12.968$, $p < 0.001$). Orthographic relatedness elicited the greatest reduction in N400 amplitude in the right hemisphere (orthographic relatedness x hemisphere interaction: medial: $F(1,82) = 4.227$, $p = 0.043$; lateral: $F(1,82) = 7.649$, $p = 0.007$), specifically at frontal electrodes in the right hemisphere (orthographic relatedness x electrode x hemisphere interaction: medial: $F(4,328) = 4.292$, $p = 0.020$; lateral: $F(2,164) = 8.234$, $p = 0.001$).

To determine if the response to words related in both orthography and phonology elicited more of a reduction in N400 amplitude than the response to words related only in orthography, a follow-up ANOVA was used to directly compare these two conditions. This analysis confirmed that words related in both orthography and phonology elicited a significantly greater reduction in N400 amplitude than did words related only in orthography (midline: $F(1,82) = 8.678$, $p = 0.004$; medial: $F(1,82) = 7.692$, $p = 0.007$).

As for phonology, there was a main effect of phonological relatedness (midline: $F(1,82) = 9.731$, $p = 0.003$; medial: $F(1,82) = 10.116$, $p = 0.002$; lateral: $F(1,82) = 3.480$, $p = 0.066$), but visual inspection of the ERP waveforms suggested that this main effect was driven by the large N400 reduction elicited by words related in both phonology and orthography (Figure 13), given that there was little-to-no reduction in the N400 elicited by words related only in phonology (Figure 12). A priori analysis of the response to phonologically related word pairs partially confirmed this hypothesis; across all participants, the difference between targets with unrelated primes and targets with phonologically related primes only approached or reached significance at medial electrodes (medial: $F(1,82) = 2.842$, $p = 0.096$) and frontal lateral electrodes (electrode interaction: medial: $F(4,328) = 2.985$, $p = 0.061$; lateral: $F(2,164) = 4.374$, $p = 0.028$). Thus, all participants as a whole did not show widely-distributed responses to words related only in phonology, though there were some significant responses.

These results did not suggest that there were significant differences in how the deaf and hearing groups responded to word pairs related in phonology. However, since one of the goals of this project was to investigate how deaf and hearing readers responded differently to words related in phonology, we *a priori* planned to analyze the response to phonologically-related word pairs separately for each group. Thus, we performed follow-up ANOVAs with the deaf and hearing participants in separate models (corrected alpha $\alpha = 0.025$ for each scalp region). These analyses indicated that for hearing participants, the difference between unrelated word pairs and phonologically related word pairs approached significance (midline: $F(1,41) = 4.099$, $p = 0.049$; medial: $F(1,41) = 3.775$, $p = 0.059$). For deaf participants, there was no difference between the unrelated word pairs and phonologically related word pairs (midline: $p = 0.929$; medial: $p = 0.684$; lateral: $p = 0.988$). These results show that the deaf participants did not show any change in N400 amplitude in response to words related only in phonology.

In summary, the grand mean ERP word pair data show that deaf and hearing participants both showed similar large N400 priming effects (N400 amplitude reductions) to word pairs related in semantics, orthography alone, and both orthography and phonology together. The N400 priming effect to words related in orthography and phonology was greater than the priming effect elicited by words related only in orthography. Deaf participants showed no priming response to words related in phonology. Hearing participants' response to words related in phonology was less clear, but suggests there was a small N400 amplitude reduction.

3.2.3. Individual differences analyses. As described in the methods section, effect magnitudes of the following ERP priming responses were calculated for use in individual differences analyses: a) semantic N400 priming effect, b) orthographic-only N400 priming effect, c) phonologic-only N400 priming effect, and d) orthographic and phonologic N400 priming effect. As was described in the methods, a negative effect magnitude indicates that the N400 was smaller in amplitude relative to the unrelated condition (i.e., a typical N400 reduction priming effect) and a positive effect magnitude indicates that the N400 was larger in amplitude relative to the unrelated condition.

Simple correlations between the four ERP priming effect magnitudes and participants' standardized reading comprehension scores can be seen in Table 9. For deaf participants, reading comprehension scores were not significantly correlated with any of the four ERP priming effect magnitudes. For hearing participants, the relationship between the semantic N400 priming effect and reading comprehension score nearly reached significance (P=0.052), with larger semantic N400 priming effects being associated with better reading comprehension scores.

Table 9. Correlation coefficients (*r*) for the relationship between word pair ERP priming effect magnitudes and standardized reading comprehension score.

	Deaf	**Hearing**
	Standardized reading comprehension score	Standardized reading comprehension score
Semantic N400 priming effect	0.031	-0.302^
Orthographic-only N400 priming effect	0.052	-0.125
Phonologic-only N400 priming effect	-0.064	-0.275^
Orthographic & phonologic N400 priming effect	-0.092	0.012

* $p < 0.05$; ^ $p < 0.1$

The relationship between the phonologic N400 priming effect and reading comprehension also approached significance for the hearing participants, again with a larger priming effect being associated with better reading comprehension scores.

Word-pair stimuli multiple regression models were previously described in Figure 5. For all three models used, all behavioral predictors showed the same predictive relationships as they did in the sentence stimuli models: for both deaf and hearing participants, an increase in the number of years of education completed predicted a higher reading comprehension score. For only the deaf participants, a higher score on the NTID speechreading test predicted a higher reading comprehension score. The results described below will focus on the relationships between the word pair ERP predictors and participants' reading comprehension score.

As the schematic in Figure 5 shows, the first set of word pair multiple regression models used the semantic N400 priming effect as the ERP predictor. The results of this model can be seen in Table 10. For deaf participants, the semantic N400 priming effect was not a significant predictor of reading comprehension score. For hearing participants, the semantic N400 priming effect approached significance as a predictor of reading comprehension score, with a greater semantic N400 priming effect magnitude predicting higher reading comprehension scores.

The second set of word pair multiple regression models included two ERP predictors, the orthographic-only N400 priming effect and the phonologic-only N400 priming effect. The results of this model can be seen in Table 11. For deaf participants, neither the orthographic-only N400 priming effect nor the phonologic-only N400 priming effect significantly predicted reading comprehension score. For hearing participants, the phonologic-only N400 priming effect approached significance as a predictor of reading comprehension score, with a greater phonologic-only N400 priming effect magnitude predicting higher reading comprehension score.

Table 10. Word-pair stimuli multiple regression models using the semantic N400 priming effect as the ERP predictor (Model 1).

	Deaf Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.64*	*22.12 (3,38)****	*.64*	*.61*	*22.12 (3,38)****					
Years of Education						3.73	(1.07)	0.41	3.50 **	0.001
NTID Speechreading						0.47	(0.11)	0.53	4.10 ***	<0.001
Grow Up Language						0.81	(0.99)	0.09	0.82	0.419
Block 2	*0.004*	*.44 (1,37)*	*.64*	*.60*	*16.46 (4,37)****					
Years of Education						3.74	(1.07)	0.41	3.48 **	0.001
NTID Speechreading						0.48	(0.12)	0.54	4.11 ***	<0.001
Grow Up Language						0.85	(1.00)	0.10	0.85	0.402
Semantic N400 priming						-0.54	(0.80)	-0.07	-0.67	0.510

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* p < .05, ** p < .01, *** p < .001. ^ p < 0.1

	Hearing Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.17*	*3.99 (2,39)**	*.17*	*.13*	*3.99 (2,39)**					
Years of Education						1.83	(0.72)	0.37	2.56 *	0.014
NTID Speechreading						0.14	(0.13)	0.16	1.09	0.284
Block 2	*.06*	*2.89 (1,38)^*	*.23*	*.17*	*3.75 (3,38)**					
Years of Education						1.72	(0.70)	0.35	2.45 *	0.019
NTID Speechreading						0.11	(0.13)	0.12	0.81	0.423
Semantic N400 priming						-1.77	(1.04)	-0.25	-1.70 ^	0.097

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* p < .05, ** p < .01, *** p < .001. ^ p < 0.1

Table 11. Word-pair stimuli multiple regression models using the orthographic-only N400 priming effect and the phonologic-only N400 priming effect as ERP predictors (Model 2).

	Deaf Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.64*	*22.12 (3,38)****	*.64*	*.61*	*22.12 (3,38)****					
Years of Education						3.73	(1.07)	0.41	3.50 **	0.001
NTID Speechreading						0.47	(0.11)	0.53	4.10 ***	<0.001
Grow Up Language						0.81	(0.99)	0.09	0.82	0.419
Block 2	*0.05*	*2.92 (2,36)^*	*.69*	*.64*	*15.78 (5,36)****					
Years of Education						4.00	(1.03)	0.44	3.89 ***	<0.001
NTID Speechreading						0.50	(0.11)	0.57	4.45 ***	<0.001
Grow Up Language						0.44	(0.96)	0.05	0.46	0.649
Ortho. N400 priming						1.51	(1.12)	0.15	1.36	0.183
Phono. N400 priming						1.27	(1.06)	0.13	1.20	0.237

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* $p < .05$, ** $p < .01$, *** $p < .001$. ^ $p < 0.1$

	Hearing Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.17*	*3.99 (2,39)**	*.17*	*.13*	*3.99 (2,39)**					
Years of Education						1.83	(0.72)	0.37	2.56 *	0.014
NTID Speechreading						0.14	(0.13)	0.16	1.09	0.284
Block 2	*.12*	*3.18 (2,37)^*	*.29*	*.22*	*3.81 (4,37)**					
Years of Education						2.18	(0.69)	0.45	3.15 **	0.003
NTID Speechreading						0.15	(0.14)	0.16	1.10	0.280
Ortho. N400 priming						-0.95	(0.83)	-0.18	-1.14	0.261
Phono. N400 priming						-1.49	(0.86)	-0.26	-1.74 ^	0.090

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* $p < .05$, ** $p < .01$, *** $p < .001$. ^ $p < 0.1$

Table 12. Word-pair stimuli multiple regression models using the combination orthographic+phonologic N400 priming effect as the ERP predictor (Model 3).

	Deaf Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.64*	*22.12* *(3,38)****	*.64*	*.61*	*22.12* *(3,38)****					
Years of Education						3.73	(1.07)	0.41	3.50 **	0.001
NTID Speechreading						0.47	(0.11)	0.53	4.10 ***	<0.001
Grow Up Language						0.81	(0.99)	0.09	0.82	0.419
Block 2	*0.001*	*.13* *(1,37)*	*.64*	*.60*	*16.24* *(4,37)****					
Years of Education						3.66	(1.10)	0.40	3.34 **	0.002
NTID Speechreading						0.48	(0.12)	0.55	4.00 ***	<0.001
Grow Up Language						0.75	(1.02)	0.08	0.73	0.468
Ortho+phono N400 priming						0.31	(0.88)	0.04	0.35	0.725

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* *p* < .05, ** *p* < .01, *** *p* < .001. ^ *p* < 0.1

	Hearing Participants									
	R^2_{change}	F_{change}	R^2_{total}	R^2_{Adj}	F_{total}	B	*(SE)*	*β*	*t*	*p*
Outcome: WRMT Reading Comprehension Score										
Block 1	*.17*	*3.99* *(2,39)**	*.17*	*.13*	*3.99* *(2,39)**					
Years of Education						1.83	(0.72)	0.37	2.56 *	0.014
NTID Speechreading						0.14	(0.13)	0.16	1.09	0.284
Block 2	*.005*	*.24* *(1,38)*	*.18*	*.11*	*.27* *(3,38)^*					
Years of Education						1.89	(0.73)	0.39	2.58 *	0.014
NTID Speechreading						0.15	(0.14)	0.17	1.12	0.269
Ortho+phono N400 priming						-0.38	(0.78)	-0.07	-0.49	0.627

Note. B *and β represent the unstandardized and standardized regression coefficiencts, respectively.*

* *p* < .05, ** *p* < .01, *** *p* < .001. ^ *p* < 0.1

The third and final set of word pair multiple regression models included the combination orthographic and phonologic N400 priming effect. The results of this model can be seen in Table 12. For both deaf and hearing participants, the combination orthographic and phonologic N400 priming effect was not a significant predictor of reading comprehension score.

In summary, the individual differences ERP word pair analyses show that overall, there was little relationship between any of the word pair ERP effect magnitudes and participants' reading comprehension scores. For hearing participants, the relationship between both the semantic and phonologic ERP priming effect magnitudes approached significance in their relationship with participants' reading comprehension scores. For both of those relationships, as the priming effect increased (i.e., as the N400 was reduced by a greater amount), participants' reading comprehension scores increased. However, this relationship did not reach significance.

Chaper 4. Discussion

4.1. Summary of results. In this study, we investigated the relationship between standardized reading comprehension scores and the magnitude of language-related ERP responses elicited while reading English in two groups: L1 English hearing adults and primarily non-native signing deaf adults, who were severely or profoundly deaf and became deaf before age two.

Grand mean analyses showed that, relative to well-formed sentences, deaf and hearing participants showed similarly large N400 responses to semantic violations alone in sentences. In response to grammatical violations alone in sentences, hearing participants showed a P600 relative to well-formed sentences, while deaf participants as a group did not show a P600. In response to combination semantic and grammatical violations ("double" violations), both deaf and hearing participants showed an N400 relative to well-formed sentences, but only the hearing participants also had a P600.

For the word-pair stimuli, grand mean analyses showed that deaf and hearing participants had similarly large N400 priming effects in response to word pairs related in semantics, orthography alone, and orthography and phonology together. For both deaf and hearing participants, the N400 priming effect was larger for words related in both orthography and phonology than for words related only in orthography. While less robust than the other responses, hearing participants appeared to show a small N400 priming effect in response to words related only in phonology. Deaf participants did not show any N400 priming effects in response to words related only in phonology.

Individual differences analyses allowed us to better understand the relationship between participants' standardized reading comprehension skill and the magnitude of their ERP responses. An important note is that while the average reading comprehension score of the deaf participants as a group was lower than the average score of the hearing participants, the best deaf and hearing participants had nearly equally high scores on the reading comprehension test, so we were truly able to assess how deaf and hearing readers attain equal levels of reading proficiency.

In terms of sentence-level ERP responses, for deaf participants, the magnitude of the N400 response to semantic violations alone or double violations was the best ERP predictor of reading comprehension score, with larger N400s predicting better reading scores. These relationships held (and indeed, were strongest) even when differences in language background, years of education, and standardized speechreading score were controlled for. For hearing participants, the magnitude of the P600 response to grammatical violations alone or double violations was the best ERP predictor of reading comprehension score, with larger P600s predicting better reading. Hearing participants also had a weak relationship between N400 effect magnitudes and reading comprehension score, with larger N400s predicting better reading scores.

In terms of word-level ERP responses, for deaf participants, there was little relationship between any of the word pair ERP priming effect magnitudes and participants' reading comprehension scores. For hearing participants, the relationship between both the semantic and phonologic ERP priming effect magnitudes approached significance in their relationship with participants' reading comprehension scores. For both of those relationships, as the priming effect increased (i.e., as the N400 amplitude was reduced by a greater amount), participants' reading comprehension scores increased. However, this relationship did not reach significance.

4.2. Relationship to prior ERP research. Before discussing the implications of these results, it is important to consider how our results relate to prior ERP research.

4.2.1. Prior L1 ERP research. The hearing readers in this study provide an opportunity to add to the small but growing literature on individual differences in L1 ERP-indexed language processing. Because this form of analysis is still relatively new, replication of results is crucial in order for the field to develop a reliable picture of how ERP responses vary as a function of L1 proficiency.

4.2.1.1. <u>Sentence-level ERP results.</u> Our sentence-level ERP results from hearing readers generally support and replicate what the prior work in this field as found. Similar to what was found in Pakulak and Neville (2010), in our hearing participants, better reading skill was predicted by larger P600s in response to grammatical violations in sentences. We also found a nearly significant relationship in which larger N400s in response to semantic violations in sentences also predicted better reading skill. There had been discrepancy in the field on this topic – one study (Newman et al., 2012) also found that higher L1 proficiency was associated with larger N400 responses to semantic violations, while a second study (Weber-Fox, Davis, et al., 2003) found that higher L1 proficiency was associated with smaller N400 responses to semantic violations. Our results replicate what Newman and colleagues found. The idea that higher proficiency might be associated with smaller N400s is not without merit, however, as there is evidence and discussion that higher cognitive skill is associated with more efficient neural processing. Further replication of results are needed to clarify the relationship between N400 magnitude and language proficiency.

Beyond the important replication of individual differences results, our study asked and answered additional questions. None of the aforementioned studies looked at both semantic and grammatical violations in the same group; our results indicate that P600 effect magnitude has a stronger association with reading comprehension skill than N400 effect magnitude does. This might be unexpected, given that the reading comprehension test we used did not have any explicit evaluations of grammatical knowledge. And yet, in the hearing participants, P600 effect magnitude was still a stronger predictor of reading comprehension score than N400 effect magnitude was.

4.2.1.2. Word priming ERP results. Many prior studies have found that higher L1 proficiency is associated with stronger N400 semantic priming effects (Balass et al., 2010; Connolly et al., 1995; Friedrich & Friederici, 2006; Henderson et al., 2011; Landi & Perfetti, 2007; Perfetti et al., 2005). In our hearing participants, larger semantic priming N400 effect magnitudes nearly reached significance in predicting better reading comprehension skill. Thus, while not statistically significant, our results are in line with the prior literature.

For phonological priming, the hearing participants as a group only had a small and approaching-significance N400 phonological priming effect. This is in contrast to many prior studies in which clear N400 phonological priming effects have been found (Coch et al., 2002; Grossi et al., 2001; Khateb et al., 2007; Rugg, 1984; Weber-Fox, Spencer, et al., 2003). However, in most of those prior studies, participants were performing a task directly related to phonology (i.e., judging if two words rhymed), which requires a focus on phonological information about words. In our lexical decision task, a focus on phonology was not required, likely leading to the reduced N400 phonological priming effects. Prior work with ERP priming effects has shown that the N400 priming effect for words related in a particular way is stronger when the relationship between those words is the focus of the task being performed (Bentin, Kutas, & Hillyard, 1993; Davids, van den Brink, van Turennout, & Verhoeven, 2011; Kramer & Donchin, 1987).

Of note, however, is that even though overall the hearing participants as a group did not show robust N400 phonological priming effects, there was a nearly significant relationship between the magnitude of the phonological priming N400 effect magnitude and reading proficiency, with larger priming effects predicting higher reading comprehension scores. This is an excellent example of an instance where averaging across a group of participants leads to a loss in the ability to see systematic differences in ERP responses.

4.2.2. Prior L2 ERP research in hearing readers. The results from the deaf participants in this study can be compared to prior work investigating individual differences in ERP responses of hearing second language (L2) learners or users. This allows us to better understand if non-native signing deaf individuals process language in similar or different ways to hearing individuals in an L2.

Prior work in hearing L2 individuals has found associations between larger N400s in response to semantic violations in sentences and better L2 proficiency or more hours of exposure to the L2 (Newman et al., 2012; Ojima et al., 2011). The results from our deaf participants match this relationship.

However, the most interesting comparison with hearing L2 individuals is the response to grammatical violations in sentences. The majority of the L2 individual differences ERP literature has shown that as people become more proficient in their L2, spend more time studying their L2, or were exposed to their L2 earlier in life, they shower larger and/or more robust P600 effects in response to grammatical violations in sentences (McLaughlin et al., 2010; Rossi et al., 2006; Tanner et al., 2014, 2013; Weber-Fox & Neville, 1996). This was not a relationship that was present in our deaf participants. There was no relationship between the magnitude of the P600 in response to grammatical violations and the deaf participants' reading comprehension scores, despite the fact that the most proficient deaf participants had as high of reading comprehension scores as the most proficient hearing participants. This suggests that non-native signing deaf individuals are not simply another group of second language learners, but that their language learning experience, especially in terms of grammatical knowledge, is markedly different from what hearing second language learners experience.

4.2.3. Prior ERP research in deaf readers. Finally, we can consider in more detail how the results from our deaf participants relate to the small prior literature on ERP reading research in deaf individuals.

Our study is best compared to the two studies by Skotara et al., which provide information on how hearing L1, hearing L2, deaf native signers, and non-native signing deaf individuals responded when reading semantic and grammatical violations in sentences (Skotara et al., 2011, 2012). Though Skotara and colleagues did not analyze individual differences in participants' ERP responses, nor how ERP responses related to language proficiency, we can still compare grand mean results between our work on theirs. In their two studies, Skotara and colleagues found that while hearing L1 and deaf native signers showed similar P600 effects in response to grammatical violations in sentences, the hearing L2 and non-native signing deaf individuals both had smaller P600s in response to grammatical violations than the hearing L1 group. All four groups had similar N400 effects in response to semantic violations in sentences. Our grand mean results with hearing L1 and non-native signing deaf groups generally match those, although our deaf participants did not actually have a significant P600 in response to grammatical violations at all. Overall, the small prior literature on ERP reading studies in deaf adults suggests that semantic processing is much more similar between deaf and hearing readers than grammatical processing is (Neville et al., 1992; Skotara et al., 2011, 2012). Our results support this idea.

The grand mean results from the two Skotara et al. studies might suggest that non-native signing deaf individuals process a written L2 in similar ways as a hearing L2 population. As was described in the prior section, however, the individual differences analyses used in this study suggest that non-native signing deaf individuals who are the most proficient readers have different patterns of neural language processing than we would expect in more proficient hearing L2 individuals. Specifically, the most proficient non-native signing deaf individuals in this study had the largest N400 effect magnitudes in response to semantic or double violations in sentences. There was no association between larger P600 magnitude and better reading comprehension, as one would expect if non-native signing deaf individuals followed similar patterns of proficiency-based changes in ERP responses to grammatical violations in sentences as hearing L2 individuals do.

In summary, our results generally match the small literature on ERP responses during reading in deaf individuals, and importantly, expand upon those results in ways that allows us to compare how non-native signing deaf individuals attain reading proficiency in different ways that hearing L1 and L2 readers do.

4.3. Implications for deaf literacy and education. The results from this study clearly indicate that non-native signing deaf individuals do not attain reading proficiency in the same ways that hearing L1 individuals do. Though we did not include a hearing L2 population in this study, results from prior research with hearing L2 individuals suggests that non-native signing deaf individuals also do not attain reading proficiency in the same ways that hearing L2 individuals do. We also found no relationship between deaf participants' phonological priming ERP responses and how well they read. What implications do these results have in terms of ways to help deaf children learn to read better?

An important caveat is warranted. Any potential changes in reading pedagogy must first be extensively tested by educational psychologists and deaf educators in order to determine their effectiveness. It can be easy to want to apply results from basic science research directly to educational practices, but any potential instructional changes must first be thoroughly tested.

4.3.1. What are successful deaf readers doing when they read? Much of the focus on the potential importance of phonological knowledge for successful reading in deaf individuals comes from the idea that deaf and hearing individuals read in similar ways (Mayer & Trezek, 2014; Paul et al., 2009; Wang et al., 2008). Our results clearly show that non-native signing deaf individuals are not reading proficiently in the same ways that hearing readers are; thus, it does not make much sense to assume that what is important for hearing children to read well are also things that deaf children need to read well. More specifically, if the phonological hypothesis were to hold, we would expect to see that the more proficient deaf readers had larger phonological priming ERP effects. We did not find any evidence of this relationship, again suggesting that phonological information is not being used by proficient non-native signing deaf individuals while reading.

If the most proficient non-native signing deaf readers are not using phonological information when they read, what are they doing instead? Our results show that the deaf participants who had the highest reading comprehension scores had larger N400 effects in response to semantic or double violations in sentences. Combined with the fact that the deaf participants had no relationship between the magnitudes of their P600s in response to grammatical violations and their reading comprehension scores, this suggests that the most successful non-native signing deaf readers are focusing on, or are most sensitive to, semantic information, not grammatical information, in sentences.

Though there is still much to learn about the neural processes that the N400 and P600 represent, it is generally thought that these two ERP responses reflect two different streams of language processing (Kuperberg, 2007; Tanner, 2013). The N400 is associated with neural processing involved in lexical access (i.e., accessing information about word meanings) and semantic integration (Kutas & Federmeier, 2000; Kutas & Hillyard, 1980, 1984), while the P600 is generally associated with the integration or reanalysis of syntactic (grammatical) information (Kaan et al., 2000; Osterhout et al., 1994; Osterhout & Holcomb, 1992), as well as with processing information when there is conflict about thematic roles (Kim & Osterhout, 2005; Kuperberg et al., 2007; van de Meerendonk et al., 2010; van Herten et al., 2006).

The best non-native signing deaf readers appear to be relying on the processing stream involving lexical access and semantic integration, rather than the syntactic processing stream. While one might think that a focus on the syntax of sentences is important in order to read well, there is a theory of sentence processing that posits that when reading sentences, readers do not actually form detailed and complete representations of the grammatical roles played by each word in a sentence. Traditional theories of sentences processing all assume that in order to understand the meaning of a sentence, a reader must access information about individual words and then use grammatical rules to organize those words into a syntactic structure, creating a "complete, detailed, and accurate representation of the linguistic input" (Ferreira, Bailey, & Ferraro, 2002). Though there are different theories of how syntactic analysis takes place, all the traditional theories of sentences processing include a thorough and complete analysis of syntactic structure prior to reaching an understanding of the meaning of a sentence (Ferreira & Clifton, 1986; Frazier, 1978; MacDonald, Pearlmutter, & Seidenberg, 1994; Truswell, Tanenhaus, & Garnsey, 1994). However, the ‘Good-Enough’ approach to sentence comprehension, first put forth by Ferreira and colleagues in 2002, challenges this idea (Ferreira et al., 2002). This theory came out of observations of sentence comprehension behavior that did not track with the idea that complete syntactic sentence structures were being computed during reading. Researchers had seen examples in which syntactic structure of sentences was not thoroughly analyzed, but rather, analyzed just enough to complete the task associated with whatever was being read (Duffy, Henderson, & Morris, 1989). In reading some sentences, individuals seemed to rely more on the meaning of individual words in the sentence, rather than their syntactic relationships, to interpret the entire meaning of the sentence (Ferreira, 2003).

The ‘Good-Enough’ theory of sentence comprehension attempts to account for these occurrences of sentence-processing behavior. In the ‘Good-Enough’ theory, when reading sentences, a reader analyzes syntactic structure only as much as is necessary for the task at hand, e.g., responding to a question or obtaining a needed piece of information. In normal day-to-day communication and reading tasks, oftentimes the majority of the necessary information can be obtained from the meanings of the words in a sentence, without needing to determine the exact grammatical role that every word in a sentence serves (Ferreira et al., 2002; Ferreira & Patson, 2007). It is reasonable to think that the most successful deaf readers in our study may be using something akin to the ‘Good-Enough’ approach to sentence comprehension when reading. Our deaf participants do not appear to be performing much syntactic integration or reanalysis of grammatical information; those who read the best are the ones with the largest ERP responses associated with meaning in sentences.

4.3.2. Implications for reading pedagogy. Now that we know that the best non-native signing deaf individuals respond most to information about meaning in sentences, what does this mean for reading instruction for deaf children?

First, we must consider what methods of reading instruction are currently used with deaf children, and their effectiveness. The research on evidence-based practices of successful reading instruction for deaf children is quite limited, and much reading instruction for deaf children is based on personal experience, teacher preference, or quasi-experimental data (Easterbrooks & Beal-Alvarez, 2013). However, a 2010 review by Easterbrooks analyzed research from a variety of fields to determine what aspects of instructional methods had strong correlational or explicitly causal relationships with better reading outcomes for deaf children. She found five "causal factors", or aspects of reading instruction, that correlated with positive reading outcomes in deaf students. The factors were: 1) focusing on teaching higher-order thinking skills such as problem solving, prediction/inferring, and general critical thinking, 2) teaching in the language(s) and modality/modalities preferred by students, 3) making information available visually and encouraging visual learning (visual imagery, organization, or memory), 4) explicitly, not implicitly, teaching skills, and 5) scaffolding instruction, that is, temporarily providing assistance in more challenging learning tasks until students can perform the tasks without assistance and thus reach a new level of knowledge (Easterbrooks, 2010). These causal factors are quite broad, and can be used across a variety of topics and teaching methods. They provide information about what general aspects of instructional methods have been most successful with deaf children.

As was discussed in the introduction, non-native signing deaf children come from a wide range of language backgrounds. Because of this, it is unlikely that one method of reading instruction is appropriate for all deaf children; pedagogy should adapt based on each child's needs (Easterbrooks & Beal-Alvarez, 2013). Still, we can use the current evidence about successful teaching strategies, combined with our results, to make suggestions about what may be more effective teaching strategies, regardless of students' language backgrounds.

Our results show that the best non-native signing deaf readers respond the most to semantic (meaning) information in sentences. If information about meaning is what is allowing successful deaf readers to read well, what are instructional methods that might help develop this knowledge and skill? Clearly, there should be a strong focus on helping children increase their vocabulary. This idea is supported by prior behavioral research showing that a strong predictor of reading achievement in deaf and hard of hearing students was productive vocabulary (in a sign or spoken language) (Kyle & Harris, 2006).

What helps deaf children learn the meanings of more words? It is important to note that there are multiple types of vocabulary knowledge. Receptive vocabulary is the vocabulary knowledge a child has in understanding a language, productive vocabulary in the vocabulary a child uses when producing a language, and reading vocabulary is the vocabulary that a child knows when reading a language (Easterbrooks & Beal-Alvarez, 2013). Much research has shown, in both hearing and deaf children, that larger receptive vocabulary is associated with a larger reading vocabulary and better overall reading proficiency (Connor, Craig, Raudenbush, Heavner, & Zwolan, 2006; Cunningham & Stanovich, 1997; Farkas & Beron, 2004; Hermans et al., 2008; Penno, Wilkinson, & Moore, 2002; Senechal, LeFebre, Thomas, & Daley, 1998). Deaf children often fall behind their hearing peers in receptive vocabulary, especially since many of them do not have as early access to language as hearing children do (Easterbrooks & Beal-Alvarez, 2013; Nicholas & Geers, 2003). Since deaf children will most easily learn a natural sign language, the most logical (but unfortunately challenging to implement) solution would be to have all deaf children raised in environments where they are surrounded by proficient sign language users. For all the deaf children who are not raised as native signers, though their receptive vocabulary development may be more challenging to develop, an emphasis on improving receptive vocabulary in their preferred method of communication is crucial (Easterbrooks & Beal-Alvarez, 2013).

There are many techniques that may help non-native signing deaf children improve their receptive vocabulary. Receptive vocabulary increases when adults have more frequent day-to-day conversations with deaf children (Easterbrooks & Beal-Alvarez, 2013). Vocabulary knowledge improves when explicit instruction in vocabulary is frequent (daily) and varied, and involves words, sentences, pictures of vocabulary concepts, and exposure to words in printed as well as signed and/or spoken forms (DeVilliers & Pomerantz, 1992; Easterbrooks & Beal-Alvarez, 2013; Paatsch, Blamey, Sarant, & Bow, 2006; Robbins & Hatcher, 1981). Hearing parents who are not fluent signers but have learned a sign language and sign to their child can help improve their child’s receptive vocabulary by signing a higher proportion of the words they say, using a larger number of signs, making sure they have their child’s attention before beginning to sign, and repeating signs many times (Easterbrooks & Beal-Alvarez, 2013; Lederberg & Everhard, 1998; Nicholas & Geers, 2003; P. Spencer, 1993).

A number of specific classroom instruction methods have either shown effectiveness in improving deaf children’s vocabulary knowledge, or have been successful for hearing children and fit within Easterbrooks’ “causal factors” - aspects of reading instruction that correlated with positive reading outcomes in deaf students. These methods include relating new words to personal experiences of students (Paul & Gustavson, 1991; Paul, 1996), relating new words to already known words or categories of words (Cannon & Easterbrooks, 2010; Easterbrooks & Beal-Alvarez, 2013), using pictures and photographs (Easterbrooks & Beal-Alvarez, 2013; Tarulli, 1998), having an abundance of words and printed information in the classroom (Easterbrooks & Beal-Alvarez, 2013; Wolfersberger, Reutzel, Sudweeks, & Fawson, 2004), and storybook reading, repetition of stories, and asking questions about stories that have been read, all in whatever method of communication the students prefer (Biemiller & Boote, 2006; Easterbrooks & Beal-Alvarez, 2013; Fung, Chow, & McBride-Change, 2005; Tavil & Soylemez, 2008). These techniques and practices have either been shown to help improve deaf children's receptive vocabulary, or show promise in doing so, and should be emphasized and tested for further understanding of their efficacy.

In addition to what we have described, there are many instructional methods that are likely important for successful reading, especially strategies that encourage more frequent reading, increase motivation to read, and involving parents in reading with children. Broader-level factors such as these are associated with better reading outcomes in deaf children (Easterbrooks & Beal-Alvarez, 2013), and are relevant independent of the specific instructional methods used to teach reading skills.

The teaching methods described above encompass many skills that are often used during reading instruction. How do our results suggest changes to what is already happening during reading instruction? Importantly, they suggest that in order to help non-native signing deaf children become proficient readers, a focus on teaching and encouraging vocabulary development may be more important than teaching grammatical information is great detail. This strategy, of course, requires further research and testing to determine its effectiveness. Despite that, our results give deaf educators and educational psychologists information about a promising direction to pursue.

4.3.3. Interactions between language background and patterns of language processing. A crucial note to make about the results of our study is that for the deaf participants, the relationships between ERP responses and standardized reading comprehension held even when differences in language background were controlled for. That is, independent of whether participants grew up exposed to only spoken language, only exposed to a form of Manually Coded English, or exposed to a mixture of both, across all participants we saw the same relationships between ERP responses and reading skill – that larger N400 responses to semantic and double violations in sentences predicted better standardized reading comprehension scores. This fact means that the potential implications for how literacy education for deaf children might be modified applied to non-native signing deaf children from any and all language backgrounds.

There is often concern that if a deaf child is not exposed to a sigh language rich environment from birth, there is little hope that she will become a proficient reader. Our results show that non-native signing deaf individuals can potentially learn to read well independent of the language background they grow up in, and that despite the wide variety of language experiences non-native signing deaf individuals have, the same patterns of neural responses to language are associated with better reading skill. This is in no way to undermine the importance of having early access to sign language, which ERP studies show is associated with more native-like responses to language while reading (Skotara et al., 2011, 2012). But as has been discussed, 90% of deaf children are born to hearing parents. It is likely easier to make changes to teaching methods used with deaf children than it is to try to get all deaf children raised with the same language experience. Our results show that for non-native signing deaf adults, regardless of the language background they have, similar patterns of neural language processing predict better standardized reading comprehension skill. The importance of this fact cannot be understated.

4.3.4. Limitations. There are certainly limitations to what our results can tell us. Importantly, we investigated what is happening in the brains of deaf adults, not deaf children. It is possible that the reading strategies used by deaf adults who are past the classroom learning stage are different than the reading strategies used by deaf children as they are learning to read. Thus, to truly understand what deaf children are doing when they are learning to read, it would be useful to use a similar paradigm as was used in this study in a group of deaf children as they are learning to read.

Additionally, we had few native signers in our group of deaf participants; for all intents and purposes, our deaf participants represent a group of non-native signing deaf individuals. So, while we cannot make conclusions about how native signing deaf individuals attain reading proficiency, the work by Skotara and colleagues (Skotara et al., 2011, 2012) suggests that deaf native signers process language while reading in similar ways to hearing L1 individuals.

In terms of assessing the importance of phonology during reading to the deaf participants, because we purposefully used a lexical decision task used in the word priming paradigm, the ERP phonological priming results represent the information about phonology that the deaf participants use or recognize when reading words for meaning. Had we used a phonological judgment task, which would require participants to focus on phonological information about words, the ERP phonological priming results would have more closely represented the sum total of each participant's phonological knowledge. While we purposefully used a lexical judgment task, in order to replicate a natural reading environment as closely as possible, we cannot assess if there is a relationship between deaf participants' reading comprehension scores and what their ERP phonological priming results would be with a phonological judgment task.

Finally, as was previously discussed, though these results provide new and important information about what patterns of language processing are associated with better reading in non-native signing deaf individuals, any potential changes in reading pedagogy must first be thoroughly tested.

Despite these limitations, this study provides important new insights into how non-native signing deaf individuals read proficiently.

4.4. Implications for ERP research. While the major purpose of this study was to investigate how some deaf individuals read more proficiently than others, our results also have important implications for future ERP research, especially for individual differences analyses.

Since individual differences analyses with ERP data are still relatively new to the field, the results from this study provide another strong example that variation in individuals' ERP responses are systematic. While individual differences analyses of ERP data are gaining more widespread support and acceptance, the results from this study show relationships of unprecedented strength between individuals' ERP responses and a measure of language proficiency (standardized reading comprehension skill, in this case). This will aid in making individual differences analyses more accepted and understood in the field. It is also important to note how useful it is to be able to relate particular ERP responses to variation in a tangible behavioral measure in standardized reading comprehension skill.

The results from the hearing participants in this study help to replicate and clarify the relationships we see between individuals' ERP responses and their reading ability. This is useful for two reasons. First, it gives the field a better picture of the neural processing mechanisms used by the most proficient hearing L1 readers of a language, which is a question many researchers want to better understand. Second, and potentially more importantly, knowing what "typically" happens in an L1 population is crucial in order for researchers to properly interpret what happens in L2 populations. Decades of grand mean ERP analyses in L1 populations have been used to interpret what ERP responses from L2 users suggest about how L2 language learning happens. Now that individual differences analyses are becoming more common, the field must again develop a robust repository of L1 individual differences ERP data, which can be used to better interpret L2 individual differences ERP data.

The results from the group of deaf participants also have great implications for the future of ERP individual differences analyses. In any group of deaf individuals, there is far more variation in language exposure, use, and proficiency than in any hearing L1 or even hearing L2 population. There could be concern that given such variation in a deaf population, it would not be possible to detect systematic relationships between ERP responses and standardized reading comprehension, due to potential co-variation with other behavioral outcome measures. However, as our results clearly show, there were clear, strong relationships between the deaf participants' reading comprehension scores and their ERP responses. These relationships were so strong that they were present even before we controlled for the effects other variables, such as speechreading score, have on reading comprehension skill. Additionally, once we controlled for the effects these variables had on reading comprehension score, the relationships between reading comprehension and ERP variables became even clearer. In short, individual differences analyses are not only possible, but even highly successful, in groups with large amounts of variability.

4.5. Final conclusions. In summary, this study showed that for equally proficient hearing L1 and non-native signing deaf adults, different patterns of neural language processing predicted better standardized reading comprehension skill, suggesting that these two populations read proficiently in different ways. Additionally, based on comparisons of prior research, our results suggest that the most proficient non-native signing deaf individuals do not have similar patterns of language processing when reading as more proficient hearing L2 individuals do.

The best non-native signing deaf readers had the largest responses to information about meaning in sentences, while the best hearing L1 readers had the largest responses to information about grammar in sentences. The results from the non-native signing deaf participants were independent of the language background in which they grew up. These results suggest that for non-native signing deaf individuals, the most important aspect of successful reading instruction may be increasing vocabulary knowledge, rather than a detailed knowledge of grammar.

In addition to the importance of these results for better understanding issues of deaf literacy, the results also have important implications for future ERP research. Our results provide clear, strong support of the feasibility of individual differences analysis of ERP data, especially in highly variable populations. The strong relationships we found between individuals' reading comprehension skill and ERP responses will be important for future decades of ERP and reading research.

References

Allen, T. (1986). Patterns of academic achievement among hearing-impaired students. In A. Schildroth & M. Karchmer (Eds.), *Deaf children in America* (pp. 161–206). San Diego, CA: College Hill Press.

Allen, T., Clark, M. D., del Giudice, A., Koo, D., Lieberman, A., Mayberry, R., & Miller, P. (2009). Phonology and reading: a response to Wang, Trezek, Luckner, and Paul. *American Annals of the Deaf*, *154*(4), 338–345. doi:10.1353/aad.0.0109

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Philadelphia: Linguistic Data Consortium.

Balass, M., Nelson, J. R., & Perfetti, C. (2010). Word learning: An ERP investigation of word experience effects on recognition and word processing. *Contemporary Educational Psychology*, *35*(2), 126–140. doi:10.1016/j.cedpsych.2010.04.001

Bellugi, U., & Studdert-Kennedy, M. (Eds.). (1980). *Signed and spoken language: Biological constraints on linguistic form.* Deerfield Beach, FL: Verlag Chemie.

Bentin, S., Kutas, M., & Hillyard, S. (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Pschophysiology*, *30*(2), 161–169.

Biemiller, A., & Boote, C. (2006). An effective method of building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*(1), 44–62.

Cannon, F., & Easterbrooks, S. (2010). Vocabulary instruction through books read in American Sign Language for English-language learners with hearing loss. *Communication Disorders Quarterly*, *31*(2), 98–112.

Capek, C. M., Grossi, G., Newman, A. J., McBurney, S. L., Corina, D., Roeder, B., & Neville, H. J. (2009). Brain systems mediating semantic and syntactic processing in deaf native signers: biological invariance and modality specificity. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 8784–8789. doi:10.1073/pnas.0809609106

Chamberlain, C., & Mayberry, R. (2008). American Sign Language syntactic and narrative comprehension in skilled and less skilled readers: Bilingual and bimodal evidence for the linguistic basis of reading. *Applied Psycholinguistics*, *29*, 367–388.

Coch, D., Grossi, G., Coffey-Corina, S., Holcomb, P. J., & Neville, H. J. (2002). A developmental investigation of ERP auditory rhyming effects. *Developmental Science*, *5*(4), 467–489. doi:10.1111/1467-7687.00241

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. doi:10.1037/0033-2909.112.1.155

Colin, S., Magnan, A., Ecalle, J., & Leybaert, J. (2007). Relation between deaf children's phonological skills in kindergarten and word recognition performance in first grade. *Journal of Child Psychology and Psychiatry*, *48*(2), 139–146.

Connolly, F., Byrne, J., & Dywan, C. (1995). Assessing adult receptive vocabulary with event-related potentials: An investigation of cross-modal and cross-form priming. *Journal of Clinical and Experimental Neuropsychology*, *17*(4), 548–565.

Connor, C., Craig, H., Raudenbush, S., Heavner, K., & Zwolan, T. (2006). The age at which young deaf children receive cochlear imlants and their vocabulary and speech-production growth: Is their an added value for early implantation? *Ear & Hearing*, *27*(6), 628–644.

Cunningham, A., & Stanovich, K. (1997). Early reading acquisition and its releation to reading experience and ability 10 years later. *Developmental Psychology*, *33*, 934–945.

Davids, N., van den Brink, D., van Turennout, M., & Verhoeven, L. (2011). Task-related influences on neurophysiological assessment of auditory rhyme: Implications for clinical studies. *Clinical Neurophysiology*, *122*(8), 1629–1636. doi:10.1016/j.clinph.2010.12.058

DeVilliers, P., & Pomerantz, S. (1992). Hearing-impaired students learning new words from written context. *Applied Psycholinguistics*, *13*, 409–431.

Duffy, S., Henderson, J., & Morris, R. (1989). Semantic facilitation of lexical access during sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 791–801.

Easterbrooks, S. R. (2010). Evidence-based curricula and practices that support development of reading skills. In M. Marschark & P. Spencer (Eds.), *The Oxford Handbook of Deaf Studies, Language, and Education* (pp. 111–126). New York NY: Oxford University Press.

Easterbrooks, S. R., & Beal-Alvarez, J. (2013). *Literacy Instruction for Students Who Are Deaf and Hard of Hearing*. New York: Oxford University Press.

Easterbrooks, S. R., & Huston, S. G. (2008). The signed reading fluency of students who are deaf/hard of hearing. *Journal of Deaf Studies and Deaf Education*, *13*(1), 37–54. doi:10.1093/deafed/enm030

Emmorey, K. (2001). *Language, Cognition, and the Brain: Insights from Sign Language Research*. Psychology Press.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, *3*(3), 464–497.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15. doi:10.1111/1467-8721.00158

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Ferreira, F., & Patson, N. D. (2007). The "Good Enough" Approach to Language Comprehension. *Language and Linguistics Compass*, *1*(1-2), 71–83. doi:10.1111/j.1749-818X.2007.00007.x

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2 : Electrophysiological evidence of the effect of L1 – L2 syntactic similarity. *Bilingualism: Language and Cognition*, *14*, 379–399. doi:10.1017/S136672891000012X

Foucart, A., & Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language*, *66*(1), 226–248. doi:10.1016/j.jml.2011.07.007

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. University of Connecticut.

Friedmann, N., & Szterman, R. (2011). The comprehension and production of Wh-questions in deaf and hard-of-hearing children. *Journal of Deaf Studies and Deaf Education*, *16*, 212–235.

Friedrich, M., & Friederici, A. D. (2006). Early N400 development and later language acquisition. *Psychophysiology*, *43*, 1–12. doi:10.1111/j.1469-8986.2006.00381.x

Fung, P., Chow, B., & McBride-Change, C. (2005). The impact of a dialogic reading program on deaf and hard-of-hearing kindergarten and early primary school-adged students in Hong Kong. *Journal of Deaf Studies and Deaf Education*, *10*(1), 82–95.

Goldin-Meadow, S., & Mayberry, R. I. (2001). How Do Profoundly Deaf Children Learn to Read? *Learning Disabilities Research & Practice*, *16*(4), 222–229. doi:10.1111/0938-8982.00022

Grossi, G., Coch, D., Coffey-Corina, S., Holcomb, P. J., & Neville, H. J. (2001). Phonological processing in visual rhyming: a developmental ERP study. *Journal of Cognitive Neuroscience*, *13*(5), 610–625. doi:10.1162/089892901750363190

Hammill, D., Brown, V., Larsen, S., & Wiederhold, J. (1994). *Test of Adolescent and Adult Language, third edition (TOAL-3)* (3rd ed.). Austin, TX: Pro-Ed.

Hänel-Faulhaber, B., Skotara, N., Kügow, M., Salden, U., Bottari, D., & Röder, B. (2014). ERP correlates of German Sign Language processing in deaf native signers. *BMC Neuroscience*, *15*(1), 62. doi:10.1186/1471-2202-15-62

Hanson, V. (1989). Phonology and reading: Evidence from profoundly deaf readers. In D. Shankweiler & I. Liberman (Eds.), *Phonology and reading disability: Solving the reading puzzle* (pp. 69–89). Ann Arbor: University of Michigan Press.

Hanson, V., & Fowler, C. (1987). Phonological coding in word reading: Evidence from hearing and deaf readers. *Memory & Cognition*, *15*, 199–207.

Henderson, L. M., Baseler, H. a., Clarke, P. J., Watson, S., & Snowling, M. J. (2011). The N400 effect in children: Relationships with comprehension, vocabulary and decoding. *Brain and Language*, *117*(2), 88–99. doi:10.1016/j.bandl.2010.12.003

Hermans, D., Knoors, H., Ormel, E., & Verhoeven, L. (2008). Modeling reading vocabulary learning in deaf children in bilingual education programs. *Journal of Deaf Studies and Deaf Education*, *13*(1), 1–20.

Jasper, H. H. (1958). The ten–twenty system of the International Federation. *Electroencephalography and Clinical Neurophysiology*, *10*, 371–375.

Jednorog, K., Marchewka, A., Tacikowski, P., & Grabowska, A. (2010). Implicit phonological and semantic processing in children with developmental dyslexia: Evidence from event-related potentials. *Neuropsychologia*, *48*, 2447–2457.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15*, 159–201.

Khateb, A., Pegna, A. J., Landis, T., Michel, C. M., Brunet, D., Seghier, M. L., & Annoni, J. M. (2007). Rhyme processing in the brain: An ERP mapping study. *International Journal of Psychophysiology*, *63*, 240–250. doi:10.1016/j.ijpsycho.2006.11.001

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*, 205–225.

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, *30*(5 SRC - GoogleScholar FG - 0), 580–602.

Klima, E., & Bellugi, U. (1979). *The Signs of Language*. Cambridge, MA: Harvard University Press.

Kotz, S., & Elston-Güttler, K. (2004). The role of proficiency on processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics*, *17*, 215–235. doi:10.1016/S0911-6044(03)00058-7

Kramer, A., & Donchin, E. (1987). Brain potentials as indices of orthographic and phonological interaction during word matching. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *13*(1), 76–86. doi:10.1037/0278-7393.13.1.76

Kroese, J., Lotz, W., Puffer, C., & Osberger, M. (1986). Chapter 10: Academic Skills. In *Language Learning Skills of Hearing-Impaired Students.* (pp. 66–76). ASHA Monographs March (23).

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research*, *1146*, 23–49. doi:10.1016/j.brainres.2006.12.063

Kuperberg, G. R., Kreher, D., Sitnikova, T., Caplan, D., & Holcomb, P. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, *100*, 223–237.

Kutas, M., & Federmeier, K. D. (2000). Electropsysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470.

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

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

Kyle, F. E., & Harris, M. (2006). Concurrent correlates and predictors of reading and spelling achievement in deaf and hearing school children. *Journal of Deaf Studies and Deaf Education*, *11*, 273–288. doi:10.1093/deafed/enj037

Kyle, F. E., & Harris, M. (2010). Predictors of reading development in deaf children: A 3-year longitudinal study. *Journal of Experimental Child Psychology*, *107*(3), 229–243. doi:10.1016/j.jecp.2010.04.011

Landi, N., & Perfetti, C. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language*, *102*, 30–45. doi:10.1016/j.bandl.2006.11.001

Lane, H., & Grosjean, F. (1980). *Recent perspectives on American Sign Language*. Hillsdale, NJ: Erlbaum Associates.

LaSasso, C., Crain, K., & Leybaert, J. (2003). Rhyme generation in deaf students: the effect of exposure to cued speech. *Journal of Deaf Studies and Deaf Education*, *8*(3), 250–270.

Lederberg, A., & Everhard, V. (1998). Communication between deaf children and their hearing mothers: The role of language, gesture, and vocalizations. *Journal of Speech, Language, and Hearing Research*, *41*, 887–899.

Lederberg, A., Schick, B., & Spencer, P. E. (2013). Language and literacy development of deaf and hard-of-hearing children: successes and challenges. *Developmental Psychology*, *49*(1), 15–30.

Luetke-Stahlman, B., & Nielsen, D. C. (2003). The contribution of phonological awareness and receptive and expressive english to the reading ability of deaf students with varying degrees of exposure to accurate english. *Journal of Deaf Studies and Deaf Education*, *8*(4), 464–484. doi:10.1093/deafed/eng028

Lyons, J. (1991). *Natural Language and Universal Grammar*. New York: Cambridge University Press.

MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

MacSweeney, M., Goswami, U., & Neville, H. (2013). The neurobiology of rhyme judgment by deaf and hearing adults: An ERP Study. *Journal of Cognitive Neuroscience*, *25*(7), 1037–1048.

Mayberry, R. I., del Giudice, A. a., & Lieberman, A. M. (2011). Reading achievement in relation to phonological coding and awareness in deaf readers: A meta-analysis. *Journal of Deaf Studies and Deaf Education*, *16*(2), 164–188. doi:10.1093/deafed/enq049

Mayer, C., & Akamatsu, C. T. (2011). Bilingualism and literacy. In M. Marschark & P. E. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (Vol. 1, pp. 144–155). New York NY: Oxford University Press.

Mayer, C., & Trezek, B. J. (2014). Is Reading Different for Deaf Individuals?: Reexamining the Role of Phonology. *American Annals of the Deaf*, *159*(4), 359–371. doi:10.1353/aad.2014.0032

McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, *7*(7), 703–704. doi:10.1038/nn1264

McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning*, *60*, 123–150. doi:10.1111/j.1467-9922.2010.00604.x

McPherson, W. B., Ackerman, P. T., Holcomb, P. J., & Dykman, R. a. (1998). Event-related brain potentials elicited during phonological processing differentiate subgroups of reading disabled adolescents. *Brain and Language*, *62*, 163–185. doi:10.1006/brln.1997.1893

Meadow-Orlans, K. P., Spencer, P. E., & Koester, L. S. (2004). *The world of the deaf infant: A longitudinal study.* USA SRC - GoogleScholar FG - 0: Oxford University Press.

Meier, R. P., & Newport, E. L. (1990). Out of the hands of babes: On a possible sign advantage in language acquisition. *Language*, *66 SRC* - , 1–23.

Mohammed, T., Campbell, R., MacSweeney, M., Barry, F., & Coleman, M. (2006). Speechreading and its association with reading among deaf, hearing and dyslexic individuals. *Clinical Linguistics & Phonetics*, *20*(7-8), 621–630. doi:10.1080/02699200500266745

Moores, D. F., & Sweet, C. A. (1990). Reading and writing skills in deaf adolescents. *International Journal of Rehabilitation Research.*, *13*(2), 178–179.

Nakano, H., Saron, C., & Swaab, T. Y. (2010). Speech and span: working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. *Journal of Cognitive Neuroscience*, *22*(12), 2886–2898. doi:10.1162/jocn.2009.21400

Neville, H. J., Kutas, M., & Schmidt, A. (1984). Event-related potential studies of cerebral specialization during reading: A comparison of normally hearing and congenitally deaf adults. *Annals of the New York Academy of Sciences*, *425*(370-376).

Neville, H. J., Mills, D., & Lawson, D. (1992). Fractionating Language: Different Neural Subsystems with Different Sensitive Periods. *Cerebral Cortex*, *2*, 244–258.

Newman, A. J., Tremblay, A., Nichols, E. S., Neville, H. J., & Ullman, M. T. (2012). The Influence of Language Proficiency on Lexical Semantic Processing in Native and Late Learners of English. *Journal of Cognitive Neuroscience*, *24*(5), 1205–1223. doi:10.1162/jocn_a_00143

Nicholas, J., & Geers, A. (2003). Hearing status, language modality, and young children's communicative and linguistic behavior. *Journal of Deaf Studies and Deaf Education*, *8*(4), 422–437.

Ojima, S., Matsuba-Kurita, H., Nakamura, N., Hoshino, T., & Hagiwara, H. (2011). Age and amount of exposure to a foreign language during childhood: Behavioral and ERP data on the semantic comprehension of spoken English by Japanese children. *Neuroscience Research*, *70*(2), 197–205.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. doi:http://dx.doi.org/10.1016/0028-3932(71)90067-4

Osterhout, L. (1997). On the brain response to syntactic anomalies: manipulations of word position and word class reveal individual differences. *Brain and Language*, *59*, 494–522. doi:10.1006/brln.1997.1793

Osterhout, L., & Holcomb, P. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*(6), 785–806. doi:http://dx.doi.org/10.1016/0749-596X(92)90039-Z

Osterhout, L., Holcomb, P. J., & Swinney, D. A. (1994). Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 786–803. doi:10.1037/0278-7393.20.4.786

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). *Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning*. (M. Carreiras & C. Clifton, Eds.)*The on-line study of sentence comprehension: Eyetracking, ERP, and beyond.* Psychology Press.

Osterhout, L., McLaughlin, J., Kim, A., Greewald, R., & Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking-ERPs, and beyond* (pp. 271–308). New York: Psychology Press.

Osterhout, L., & Nicol, J. (1999). On the Distinctiveness, Independence, and Time Course of the Brain Responses to Syntactic and Semantic Anomalies. *Language and Cognitive Processes*, *14*(3), 283–317. doi:10.1080/016909699386310

Paatsch, L., Blamey, P., Sarant, J., & Bow, C. (2006). The effects of speech production and vocabulary training on different components of spoken language performance. *Journal of Deaf Studies and Deaf Education*, *11*(1), 39–55.

Padden, C., & Ramsay, C. (2000). American Sign Language and reading ability in deaf children. In C. Chamberlain, J. Mordord, & R. Mayberry (Eds.), *Language acquisition by eye* (pp. 143–163). Mahwah, NJ: Earlbaum.

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, *22*(12), 2728–2744. doi:10.1162/jocn.2009.21393

Paul, P. (1996). Reading vocabulary knowledge and deafness. *Journal of Deaf Studies and Deaf Education*, *1*(1), 3–15.

Paul, P., & Gustavson, G. (1991). Hearing-impaired students' comprehension of high-frequency multimeaning words. *Remedial and Special Education (RASE)*, *12*(4), 52–62.

Paul, P., Wang, Y., Trezek, B. J., & Luckner, J. L. (2009). Phonology is necessary, but not sufficient: a rejoinder. *American Annals of the Deaf*, *154*(4), 346–356. doi:10.1353/aad.0.0110

Penno, J., Wilkinson, I., & Moore, D. (2002). Vocabulary acquisition from teaching explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *94*, 23–33.

Perfetti, C. (1985). *Reading ability.* New York, NY: Oxford University Press.

Perfetti, C., & Sandak, R. (2000). Reading optimally builds on spoken language: implications for deaf readers. *Journal of Deaf Studies and Deaf Education*, *5*(1), 32–50. doi:10.1093/deafed/5.1.32

Perfetti, C., Wlotko, E. W., & Hart, L. a. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *31*(6), 1281–1292. doi:10.1037/0278-7393.31.6.1281

Phillips, N. A., Segalowitz, N., Brien, I. O., & Yamasaki, N. (2004). Semantic priming in a first and second language : evidence from reaction time variability and event-related brain potentials. *Journal of Neurolinguistics*, *17*, 237–262. doi:10.1016/S0911-6044(03)00055-1

Prat, C. S., & Just, M. A. (2011). Exploring the neural dynamics underpinning individual differences in sentence comprehension. *Cerebral Cortex*, *21*(8), 1747–1760.

Qi, S., & Mitchell, R. E. (2012). Large-Scale Academic Achievement Testing of Deaf and Hard-of-Hearing Students : Past , Present , and Future. *Journal of Deaf Studies and Deaf Education*, *17*(1), 1–18. doi:10.1093/deafed/enr028

Rayner, K., Foorman, B. R., Perfetti, C., Pesetsky, D., & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, *2*, 31–74.

Reilly, J., & McIntire, M. (1980). American Sign Language and Pidgin Sign English: What's the Difference? *Sign Language Studies*, *27*, 151–192.

Robbins, N., & Hatcher, C. (1981). The effects of syntax on the reaing comprehension of hearing impaired children. *The Volta Review*, *83*, 105–115.

Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The Impact of Proficiency on Syntactic Second-language Processing of German and Italian: Evidence from Event-related Potentials. *Journal of Cognitive Neuroscience*, *18*(2), 2030–2048.

Rugg, D. (1984). Event-Related Potentials in Phonological Matching Tasks. *Brain and Language*, *23*, 225–240.

Sandler, W., & Lillo-Martin, D. (2001). Natural sign languages. In M. Aronoff & J. Rees-Miller (Eds.), *Hanbook of Linguistics* (pp. 533–562). Malden, MA: Blackwell Publishers.

Senechal, M., LeFebre, J., Thomas, E., & Daley, K. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly*, *32*, 96–116.

Sims, D. (2009). *NTID Speechreading: CID Everyday Sentences Test.* Rochester, NY: Rochester Institute of Technology.

Skotara, N., Kügow, M., Salden, U., Hänel-Faulhaber, B., & Röder, B. (2011). ERP correlates of intramodal and crossmodal L2 acquisition. *BMC Neuroscience*, *12*(48). doi:10.1186/1471-2202-12-48

Skotara, N., Salden, U., Kügow, M., Hänel-Faulhaber, B., & Röder, B. (2012). The influence of language deprivation in early childhood on L2 processing : An ERP comparison of deaf native signers and deaf signers with a delayed language acquisition. *BMC Neuroscience*, *13*(44). doi:10.1186/1471-2202-13-44

Spencer, L., Barker, B., & Tomblin, J. (2003). Exploring the Language and Literacy Outcomes of Pediatric Cochlear Implant Users. *Ear & Hearing*, *24*(3), 236–247. doi:10.1097/01.AUD.0000069231.72244.94

Spencer, P. (1993). Communication behaviors of infants with hearing loss and their hearing mothers. *Journal of Speech and Hearing Research*, (36), 311–321.

Stanovich, K., & West, R. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, *24*, 402–433.

Stocco, A., & Prat, C. (2014). Bilingualism trains specific brain circuits involved in flexible rule selection and application. *Brain and Language*, *137*, 50–61.

Stokoe, W., Casterline, D., & Croneberg, C. (1965). *A dictionary of American Sign Language on linguistic principles*. Washington, DC: Gallaudet University Press.

Strong, M., & Prinz, P. (2000). Is American Sign Language skill related to English literacy? In C. Chamberlain, J. Mordord, & R. Mayberry (Eds.), *Language acquisition by eye* (pp. 143–163). Mahwah, NJ: Earlbaum.

Supalla, S. (1991). Manually Coded English: The Modality Question in Sign Language Development. In P. Siple & S. Fischer (Eds.), *Theoretical Issues in Sign Language Research Volume 2: Psychology* (pp. 85–109). Chicago: University of Chicago Press.

Tanner, D. (2013). Individual differences and streams of processing. *Linguistic Approaches to Bilingualism*, *3*(3), 350–356.

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in on-line L2 grammatical comprehension. *Bilingualism: Language and Cognition*, *17*(2), 277–293. doi:10.1017/S1366728913000370

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, *16*(2), 367–382. doi:10.1017/S1366728912000302

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289–301. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2014.02.002

Tarulli, N. (1998). Using photography to enhance language and learning: A picture can encourage a thousand words. *Language, Speech, and Hearing Services in Schools*, *29*, 54–57.

Tavil, Z., & Soylemez, A. (2008). Vocabulary teaching through storytelling to very young learners in kindergarten. *Ekev Academic Review*, *12*(35), 371–382.

Traxler, C. B. (2000). The Stanford Achievement Test, 9th Edition : National Norming and Performance Standards for Deaf and Hard-of-. *Journal of Deaf Studies and Deaf Education*, *5*, 337–348.

Treiman, R., & Hirsh-Pasek, K. (1983). Silent reading: Insights from second-generation deaf readers. *Cognitive Psychology*, *15*, 39–65.

Truswell, J., Tanenhaus, M., & Garnsey, S. (1994). Semantic influences on parsing: Use of thematic role information in syntactic disambiguation. *Journal of Memory and Language*, *33*, 285–318.

Trybus, R. J., & Karchmer, M. A. (1977). School achievement scores of hearing impaired children: national data on achievement status and growth patterns. *American Annals of the Deaf*, *122*(2), 62–69.

Van de Meerendonk, N., Kolk, H., Vissers, C., & Chwilla, D. (2010). Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, *22*, 67–82.

Van Herten, M., Chwilla, D., & Kolk, H. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*, 1181–1197.

Wang, Y., Trezek, B. J., Luckner, J. L., & Paul, P. V. (2008). The Role of Phonology and Phonologically Related Skills in Reading Instruction for Students Who Are Deaf or Hard of Hearing. *American Annals of the Deaf*, *153*(4), 396–407. doi:10.1353/aad.0.0061

Wang, Y., & Williams, C. (2014). Are We Hammering Square Pegs Into Round Holes?: An Investigation of the Meta-Analyses of Reading Research With Students Who Are d/Deaf or Hard of Hearing and Students Who Are Hearing. *American Annals of the Deaf*, *159*(4), 323–345. doi:10.1353/aad.2014.0029

Weber-Fox, C., Davis, L. J., & Cuadrado, E. (2003). Event-related brain potential markers of high-language proficiency in adults. *Brain and Language*, *85*, 231–244. doi:10.1016/S0093-934X(02)00587-4

Weber-Fox, C., & Neville, H. J. (1996). Maturational Constraints on Functional Specializations for Language Processing : ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256.

Weber-Fox, C., Spencer, R., Cuadrado, E., & Smith, A. (2003). Development of neural processes mediating rhyme judgments: Phonological and orthographic interactions. *Developmental Psychobiology*, *43*(2), 128–145.

Wolfersberger, M., Reutzel, D., Sudweeks, R., & Fawson, P. (2004). Developing and validating the classroom literacy environment profile (CLEP): A tool for examining the “print richness” of early childhood and elementary classrooms. *Journal of Literacy Research*, *36*(2), 211–272.

Woodcock, R. (2011). *Woodcock Reading Mastery Tests, Third Edition*. Bloomington, MN: Pearson.

Woodward, J. (1973). Some Characteristics of Pidgin Sign English. *Sign Language Studies*, *3*, 39–46.