Language and Linguistics Compass 9/10 (2015): 406–419, 10.1111/lnc3.12167

Individual Differences in Language Processing: Electrophysiological Approaches

Megan A. Boudewyn*
University of California, Davis

Abstract

Language processing is a complex task that requires both specialized cognitive processes (e.g. speech decoding) and more general cognitive processes (e.g. working memory). Research on how individual differences in these processes influence language processing and comprehension has primarily relied on behavioral methods, such as reaction time measures, self-paced reading, and eye-tracking. However, a growing number of studies have used electrophysiological (EEG) techniques to study individual differences in language processing. EEG and event-related potential (ERP) methods provide a unique link between neural activity and cognitive processing and can be used to draw specific inferences about the neural basis of language processing and its variability. The primary goal of this paper is to showcase EEG/ERP studies that have made significant contributions to the study of individual differences in how the brain processes language, over and above what would be possible using behavioral methods alone. A secondary goal of this paper is to highlight several methodological issues specific to research on individual differences in language processing and identify ways in which EEG/ERP studies can take advantage of what has been learned from previous research to minimize these issues.

Introduction

Language processing is a complex cognitive task, one that requires specialized processes, such as those involved in the decoding of words from a speech stream, and more general cognitive processes, such as those involved in maintaining task-relevant information in memory. Individual differences in how language is processed arise from variability in both specialized and domain-general processes and from variability in factors such as language experience, age, interest, and motivation. Understanding how these sources of variability contribute to individual differences in language processing provides insight into how the language system works and informs our understanding of language disorders and difficulties.

My primary goal in this paper is to showcase electrophysiological studies that have made unique and significant contributions to our understanding of variability in how the brain processes language, over and above what would be possible using behavioral methods alone. A secondary goal of this paper is to highlight several methodological issues specific to research on individual differences in language processing and identify ways in which electrophysiological studies can take advantage of what has been learned from previous research to minimize these issues.

Section 1. Individual Differences in Language Processing: A Short Background

There is a rich literature on individual differences in language processing and comprehension. A thorough discussion of this literature is beyond the scope of the current review. Instead, my goal in this section is to provide a brief overview of major sources of variability in language processing in order to ground the discussion of what ERP research in this area has to offer.

LANGUAGE PROCESSING DEPENDS ON SPECIALIZED PROCESSES

A number of specialized cognitive processes are engaged during language comprehension. For example, in order to understand a word, an individual must first identify that word's form. In speech, that involves identifying the phonological form from the speech stream; in reading, it involves mapping an orthographic form to a phonological one. Although these are relatively 'low-level' processes involved in the identification of single words, individual differences in their execution are substantial. In reading, individual differences in the quality of lexical representations affect the speed and accuracy of word decoding, which in turn contribute to variability in sentence- and discourse-level processing. The relation between word-decoding ability and comprehension is described in a framework called the Lexical Quality Hypothesis (Perfetti, 2007; Perfetti & Hart, 2002). According to the hypothesis, lexical representations are developed over time through repeated encounters with a given word and serve to connect the word's phonological, orthographic, and semantic features. Poor-quality lexical representations are argued to negatively impact comprehension by consuming resources that would otherwise be allocated to higher-order processing at the sentence and discourse levels. In addition, poor-quality lexical representations place limits on how fully words are understood, meaning that relevant background knowledge might not be activated appropriately (Perfetti, 2007; Perfetti & Hart, 2002).

Several studies have found that performance on word-decoding tasks is predictive of comprehension performance, even in skilled adult readers (Bell & Perfetti, 1994; Hamilton, Freed, & Long, 2013; Perfetti, Wlotko, & Hart, 2005). These tasks generally require individuals to make judgments about which of two letter strings is a real word, either based on how they are written (orthographic judgment) or how they would sound if pronounced (phonological judgment). Faster decision times are taken to reflect better quality lexical representations. For example, in one recent study, individuals who performed relatively poorly on orthographic and phonological judgment tasks exhibited relatively slow reading times after encountering a new noun in a text (Hamilton et al., 2013). This effect was mediated by working memory capacity, which highlights that individual differences in language processing arise from a complex combination of factors. Even so, this finding shows that 'low-level' word-decoding ability is important and influential in determining how language is processed, even among skilled college-level readers.

LANGUAGE PROCESSING DEPENDS ON OTHER COGNITIVE PROCESSES

In addition to specialized cognitive processes, language processing engages attention, working memory, and cognitive control processes. Of these, research on individual differences in language processing has primarily focused on the role of working memory and cognitive control. Working memory refers to a limited-capacity system responsible for the short-term maintenance, storage, and manipulation of information. Cognitive control is a term encompassing a range of executive functions, including conflict detection and resolution, goal/context maintenance, and inhibition mechanisms. These constructs have some degree of overlap, and may not be separable; certainly, tasks designed to assess working memory and cognitive control are highly correlated (Engle, 2002; Kane, Bleckley, Conway, & Engle, 2001; Kane & Engle, 2003). Regardless of whether working memory and cognitive control are theoretically or empirically separable, the subprocesses that are typically ascribed to these constructs are crucially important to a wide range of cognitive functions, including language. Namely, inhibition of irrelevant information, context/goal maintenance, error detection, and monitoring are important during language processing, no matter how one might choose to group them in relation to the broader constructs of working memory and cognitive control.

For example, when a reader encounters an ambiguous word (e.g. *bank*), the context-irrelevant meaning may be initially activated but should then be suppressed (e.g. the side-of-the-river meaning in a sentence about the financial institution). Models of discourse processing have traditionally stressed the importance of suppressing information that is automatically activated during reading, but is context-irrelevant, when constructing coherent, well-organized representations of language context (Gernsbacher, 1993, 1996, 1997; Kintsch, 1988; Kintsch & Van Dijk, 1978; Van Dijk & Kintsch, 1983). Individual differences in the suppression of context-irrelevant information have likewise featured prominently as a proposed source of variability in language processing (Gernsbacher, 1993, 1996, 1997). Several studies have provided support for this idea. In one study, both skilled and less-skilled readers successfully selected test words relating to the context-appropriate meaning of an ambiguous word (e.g. *garden* after *He dug with the spade*), but less-skilled readers were slower than skilled readers at rejecting test words that were related to the context-inappropriate meaning (e.g. *ace* after *He dug with the spade*) (Gernsbacher, Varner, & Faust, 1990). Similarly, less-skilled readers showed greater facilitation than skilled readers for context-appropriate meanings following biased contexts (e.g. *garden* after *He dug with the spade*) compared with neutral contexts (*garden* after *He picked up the spade*) (Gernsbacher & Faust, 1991).

Complementary to the inhibition of context-irrelevant information is the maintenance of relevant context. In order to fully understand incoming words and sentences, it is necessary to consider the preceding context in which they occur. Failure to maintain or rapidly retrieve relevant context would, for example, make the interpretation of pronouns quite difficult. For instance, a pronoun anaphor such as 'she' or 'he' would have limited meaning (i.e. who is she/he?) unless the reader or listener is able to link it to its antecedent. Likewise, the selection of the appropriate sense of ambiguous words requires taking preceding context into account. Without having access to relevant context, how would a reader or listener know which sense of 'bank' is intended (financial institution or side of a river), and which sense should be suppressed as context-irrelevant? It is therefore not surprising that the manner in which context is taken into account during the processing of incoming words and sentences has been a major focus of research in psycholinguistics. Similarly, individual differences in the capacity to maintain or retrieve relevant context have been a major focus of research on variability in language processing.

There are several theoretical accounts of how individual differences in working memory capacity affect language processing. One prominent view holds that limitations in working memory capacity lead to processing costs for computationally demanding language (such as complex syntactic structures), which 'takes up' most or all of working memory resources, leaving fewer resources available for the storage of contextual information (Daneman & Carpenter, 1980; Just & Carpenter, 1992; King & Just, 1991). There has been considerable debate surrounding whether syntactic computations tap into this kind of general working memory resource (Clifton et al., 2003; Traxler, Williams, Blozis, & Morris, 2005; Waters & Caplan, 1996). At the discourse level, working memory might be expected to constrain the amount of context that an individual can maintain at a given time. Another prominent view holds that during discourse comprehension, readers/listeners do not necessarily maintain large quantities of context as 'active' in working memory but instead acquire the ability to store representations of language context in long-term memory and quickly retrieve them when cued by incoming words and ideas (Ericsson & Kintsch, 1995). Similarly, recent work has suggested that individual differences in language comprehension that are commonly attributed to variability in working memory capacity might be better explained by individual differences in susceptibility to interference, or in other words, to retrieval failures rather than to maintenance failures (Van Dyke & Johns, 2012; Van Dyke, Johns, & Kukona, 2014). In any case, the ability to access previous context (through maintenance or through retrieval) would be expected to impact how individuals process incoming information during language comprehension.

LANGUAGE PROCESSING DEPENDS ON EXPERIENCE

Finally, it is important to point out that how an individual processes language depends on experience. Both the specialized processes dedicated to language processing, such as word-decoding ability, and the other cognitive processes that are engaged by language processing, such as working memory maintenance, can vary as a function of experience. For example, as discussed above, high-quality lexical representations are developed through repeated encounters with a word. Therefore, lexical representations should improve over time for an individual as language skills develop, and also as a function of frequency of use (i.e. voracious readers should have relatively high-quality lexical representations). Vocabulary also increases over time and with increased print exposure. Similarly, it has been suggested that the ability to rapidly store and retrieve large amounts of language context increases with practice and reading skill (Ericsson & Kintsch, 1995). Thus, most recent models of individual differences in language processing try to take language experience and skill into account by including measures of print exposure and vocabulary (e.g. Braze, Tabor, Shankweiler, & Mencl, 2007; Hamilton et al., 2013; Misyak & Christiansen, 2012).

Section 2. Individual Differences in Language Processing: What Does Electrophysiology Offer?

Why choose electrophysiological methods to study individual differences in language processing? Electroencephalogram (EEG) recordings measure changes in electrical voltage caused by brain activity. Activity in the EEG signal that is synchronized to a particular event type (e.g. the onset of words in experimental condition X) is known as an event-related potential (ERP). In addition to ERPs, information about oscillatory changes in various frequency bands can be extracted from the EEG signal. These methods provide information about language and cognitive processing with millisecond precision. In addition, EEG can be recorded while participants read or listen to language for comprehension, without the necessity for a secondary task (although many studies do include some kind of task). As such, EEG/ERP methods provide a unique link between neural activity and cognitive processing, and they can be used to make specific inferences about variability in how the brain processes language. In this section, I highlight studies that have taken advantage of electrophysiological indices that are reliably sensitive to specific linguistic and cognitive manipulations in order to study variability in language processing.

LEVERAGING LANGUAGE-LINKED ERP EFFECTS TO STUDY VARIABILITY IN PROCESSING

There are a number of language-linked ERP effects that have been well-characterized. Perhaps the most widely studied effect is the N400. The N400 is a negative-going waveform elicited by words (or other meaningful stimuli, such as pictures). It peaks at about 400 ms after word onset and is reduced as a function of word characteristics (e.g. smaller amplitudes for common words than for rare words) and semantic fit (e.g. smaller amplitudes for contextually supported words than for unsupported words) (for reviews, see Kutas & Federmeier, 2000; Kutas & Federmeier, 2011; Kutas & Van Petten, 1994; Kutas, Van Petten, & Kluender, 2006; Swaab, Ledoux, Camblin, & Boudewyn, 2011). In contrast, other language-linked ERPs have been shown to be sensitive to different aspects of linguistic input and are differentiated from the N400 by timing, polarity, distribution across the scalp, and task sensitivity. One well-known example is the P600. The P600 is a positive-going ERP that peaks around 600 ms post word onset and is typically found in response to syntactic difficulty or complexity, or to conflict between syntactic and semantic processing (for reviews, see Kaan, Harris, Gibson, & Holcomb, 2000; Kolk & Chwilla, 2007; Kuperberg, 2007; Swaab et al., 2011). A more nuanced description of the functional significance of each of these effects is beyond the scope of this review, but it is important to note that the traditional 'N400-semantics/P600-syntax' dichotomy is an oversimplification. Still, the presence of an N400 effect as opposed to a P600 effect in a given experimental context (or vice versa) is informative: the N400 indexes lexical/semantic processing, whereas the P600 reflects some kind of combinatorial processing, whether that be syntactic revision or event prediction error, as has been recently suggested (Kuperberg, 2013). A number of studies have fruitfully leveraged the differences in these effects to contribute to the literature on variability in language processing.

For example, the N400 and P600 have been used to investigate how individual differences in working memory influenced how listeners process sentences as they unfold during comprehension (Nakano, Saron, & Swaab, 2010). Nakano et al. had participants listen to sentences in which the plausibility of the first noun taking on a subject role was varied (e.g. *The dog/poet/box….*). Nouns in this position are most frequently assigned the thematic role of agent, meaning the 'doer' of the action that will be described in the sentence. Therefore, encountering an inanimate noun (which does not make a very good agent; e.g. *box*) in this position may serve as a signal to the listener that the sentence will contain an infrequent structure, such as one in which the first noun does not take on an agent role. ERPs that were time-locked to the first noun showed an effect of animacy, such that a larger frontal negativity was elicited by inanimate nouns (e.g. *box*, which cannot plausibly take on an agent role) compared with animate nouns (e.g. *dog/poet*, which can plausibly take on an agent role). Importantly, sensitivity to the animacy cue varied as a function of performance on WM span tasks[1]: only individuals who performed relatively well on the span tasks were sensitive to the difference between animate and inanimate nouns in this position (Nakano et al., 2010).

Of particular interest in this study is that sensitivity to the animacy cue had consequences for how the rest of the sentences were processed. After the initial noun, sentences continued on to be plausible or less plausible given world knowledge (e.g. *The dog/poet/box was biting the mailman.*). At the verb (*biting*), individuals who performed well on WM span tasks showed a P600 effect (P600 to *box…biting* compared with *dog…biting*), whereas individuals who performed relatively poorly on the WM span tasks showed an N400 for the same contrast (larger N400 to *box…biting* compared with *dog…biting*) (Nakano et al., 2010). This qualitative difference indicates that individuals who performed well on the WM span tasks used the animacy of the first noun as a signal about the likely syntactic structure/thematic role assignment of the sentence, assuming that the first noun in the inanimate condition (*box*) would turn out to have a non-agent role. This then led to conflict and the need for revision upon encountering the verb (*biting*), which requires that the first noun be assigned to an agent role, even though that was semantically implausible. In other words, it does not make sense that the *box was biting*, but the syntax dictates that this is the correct parse. In contrast, individuals who performed poorly on the WM span tasks appear to have been uninfluenced by the animacy signal at the first noun, assuming the 'default' agent role for all nouns in this position. This then resulted in the verb in the inanimate condition (*box…biting*) being processed as a semantic violation (yielding an N400 effect), but not one that caused conflict with the presumed syntactic structure (Nakano et al., 2010). This study provides an elegant example of how ERPs can be used to provide unique information about individual differences in language processing. In this case, one might expect that an analogous behavioral study in which reading times were measured might not capture these differences, as all individuals would be expected to slow down upon encountering the critical verb in the inanimate condition regardless of the nature of the processing difficulty. By using ERPs, however, Nakano et al. (2010) were able to leverage the N400 and P600 in order to provide insight into the processing difficulties encountered by individuals as a function of their performance on WM span tasks.

Another line of research has capitalized on the distinction between the N400 and P600 effects in order to study individual differences in second-language learning. In a series of studies, ERPs have been recorded as native speakers and learners of languages including French, German, and Spanish read sentences containing semantic and grammatical anomalies, in order to track variability in the internalization of grammatical rules during second-language learning (Osterhout et al., 2008; Tanner, Inoue, & Osterhout, 2014; Tanner, McLaughlin, Herschensohn, & Osterhout, 2013; Tanner, Osterhout, & Herschensohn, 2009). For example, in one study, college students taking an introductory French course read sentences containing semantic violations (e.g. *Sept plus livre font douze* [*Seven plus book make twelve*]) and verb conjugation violations (e.g. *Tu adorez le francais* [*You (singular) adore (plural) the French*]), among other violations (Osterhout et al., 2008). In native speakers, this type of semantic violation elicits an N400 effect, whereas this type of grammatical violation elicits a P600 effect. After just 1 month of French instruction, about half of the students tested showed an N400 effect in response to both violation types, compared with non-anomalous control conditions (Osterhout et al., 2008). This subset of participants (the 'fast learner' group) was also more accurate at identifying sentences containing violations than the other participants tested. Interestingly, when re-tested after 4 months of instruction, the ERP signature elicited by the grammatical violations had changed in the fast learner group, whereas grammatical violations had initially yielded an N400 effect when compared with a control condition. This comparison yielded a P600 effect after additional experience with the language. Although the sample size in this study was relatively small (N = 14; N = 7 in the fast learner group), the pattern of results has been replicated in additional studies testing students in introductory second-language courses (Tanner et al., 2014; Tanner et al., 2013; Tanner et al., 2009). Namely, during second-language learning, N400 effects in response to grammatical violations transition to P600 effects over time as experience with the language increases, with P600 effects similar to those elicited by native speakers appearing in all students tested after 3 years of instruction (Tanner et al., 2013). This suggests that second-language learners start by memorizing specific combinations of words and verb conjugations, much like a new word might be learned, before internalizing the underlying grammatical rule (Osterhout et al., 2008). This line of work provides another example of how language-linked ERPs like the N400 and P600 can provide detailed information about variability in language processing, in this case in the context of second-language learning.

Finally, at least one ERP study has used the N400 effect to investigate individual differences in word-decoding ability and comprehension (Perfetti et al., 2005). Perfetti et al. (2005) trained skilled and less-skilled readers (as measured by performance on a standardized comprehension test) on the meanings of rare, unfamiliar words (e.g. *gloaming*, meaning 'twilight'). ERPs were then recorded while participants were presented with prime-target pairs. Prime words included trained rare words, untrained rare words, and untrained familiar words; targets were either semantically related or unrelated to their primes. Results showed a larger positive deflection to the trained prime words compared with the untrained primes, interpreted as a memory effect commonly found for 'old' compared with 'new' items. Moreover, this effect was larger for skilled readers than for less-skilled ones (Perfetti et al., 2005). At the target words, N400-like effects of relatedness (larger negative deflection for unrelated targets compared with related targets) were found in the trained rare condition as well as the untrained familiar condition, but not in the untrained rare condition. Importantly, skilled readers showed larger effects than less-skilled readers (Perfetti et al., 2005). These results indicate that skilled readers more effectively learned new words, reflected in a stronger episodic memory ERP effect for untrained compared with trained prime words. Skilled readers were also better than less-skilled readers at linking new word forms to meanings, reflected in a stronger N400 effect of relatedness for trained rare words (Perfetti et al., 2005). This study provides an example of how the N400 can be used to investigate variability in word learning and word-decoding ability.

LEVERAGING ELECTROPHYSIOLOGICAL INDICES TO STUDY HOW LANGUAGE PROCESSES RELATE TO OTHER COGNITIVE PROCESSES

A number of studies have used electrophysiological indices related to cognitive processes such as working memory to study individual differences in language processing. One line of research has focused on sustained changes observable in the EEG signal at frontal electrode sites, known as 'slow waves'. Unlike the N400 or P600 effects discussed above, slow-wave effects refer to differences between conditions that are sustained over relatively long periods of time (e.g. a second or more, rather than a couple of hundred milliseconds or less). Sustained effects such as this offer unique insight into neural/cognitive processes that unfold over relatively long periods of time, like WM maintenance demands that are associated with the accumulation of language context during comprehension.

Slow-Wave Effects

Increased sustained frontal negativities have been observed in response to sentences containing syntactically complex structures (object-relatives: *The reporter that the senator attacked committed the error*) compared with syntactically simpler structures (subject-relatives: *The reporter that attacked the senator committed the error*) (King & Kutas, 1995; Müller, King, & Kutas, 1997). Similar sustained frontal effects have been observed for other comparisons in which the 'context load' demands of one condition were more demanding than the other. In one study, a larger sustained frontal negativity was found in response to unambiguous sentences when syntactically connected elements were separated (and had to be maintained) across a long portion of the sentence, compared with over a shorter portion of the sentence (Fiebach, Schlesewsky, & Friederici, 2002). A similar effect was found when comparing unambiguous sentences read while maintaining a high concurrent memory load compared with a low concurrent memory load (Vos, Gunter, Kolk, & Mulder, 2001), and with sentences containing temporarily ambiguous syntactic structures, when disambiguation was not cued by prior context compared with when prior context contained a disambiguation cue (Vos & Friederici, 2003). An increased sustained frontal negativity was also found when participants read sentences in which events were described out of sequence (e.g. before Y, X) compared with sentences in which events were described in sequential order (e.g. after X, Y) (Münte, Schiltz, & Kutas, 1998). Overall, the pattern seems to be that sustained frontal negative deflections show up when maintenance of the developing language context is relatively demanding.

Several of these studies found individual differences in the slow-wave effects, either as a function of performance on WM span tasks (Fiebach et al., 2002; Vos et al., 2001), or of performance on comprehension questions that followed the test sentences during EEG recording (King & Kutas, 1995; Müller et al., 1997). Across studies, the pattern of results is complex. Individuals who performed relatively poorly on a WM span task showed larger frontal negative shifts than individuals who performed relatively well on a WM span task, when sentences were syntactically demanding (Fiebach et al., 2002), and when a high WM load was added to the sentence-reading task (Vos et al., 2001). However, in other studies, individuals who performed relatively well on the comprehension questions that followed test sentences showed larger frontal negative shifts compared with individuals who performed poorly on the comprehension questions, when sentences were syntactically demanding (King & Kutas, 1995; Müller et al., 1997). In Münte et al. (1998), individuals who performed relatively well on a WM span task showed larger slow-wave effects than individuals who performed relatively poorly.

Overall, larger negative deflections seem to be associated with the more difficult of two conditions (whether that means sentences containing complex syntactic structures, out-of-order events, or a high concurrent memory load), but whether or not it is the 'good' or the 'poor' performers who show more robust effect sizes depends on how the performance is measured. There may be a way to reconcile these findings. In the studies in which 'poor' performance was associated with smaller slow-wave effects, performance on WM span tasks was not available, and accuracy on comprehension questions that followed test items was used to classify participants as 'good' or 'poor' performers. By definition then, individuals in the poor performance group did not understand the sentences as well as those in the good performance group. It is therefore not necessarily surprising that individuals with the best sentence understanding were the most sensitive to difference between syntactically complex and syntactically simpler sentence types. In contrast, in those studies in which performance on a separate WM span task was available, dividing participants into high WM and low WM performance resulted in larger slow-wave effects for the low WM performance group. This pattern suggests that, when readers understand the test sentences, larger slow-wave effects are associated with poor WM performance. This would support the idea of the slow frontal negative shift as an index of increased WM demands. Of course, one caveat to this account is that in the work by Münte et al. (1998), good WM performance was linked to larger slow-wave effects, although this was the only study in which the experimental manipulation involved semantic cues (the presence of *before* vs. *after* to indicate event sequence). It is possible that this pattern may have been driven by the difference in content between this study and others. In any case, these studies provide an interesting example of how electrophysiological effects associated with WM or other cognitive processes can be used to investigate variability in language processing.

Time-Frequency Approaches

In addition to ERPs and slow waves, EEG can be studied using time-frequency approaches, which focus on changes in different frequency bands over time. Time-frequency methods themselves are not a new methodological development, but they have not yet been extensively applied to the study of language processing. This literature is rapidly growing, however, and some studies have taken advantage of the link between changes in various frequency bands and cognitive processes such as working memory and attention in order to study individual differences in language processing.

For example, EEG activity has been used to study individual differences during language development that have potential clinical relevance (Benasich, Gou, Choudhury, & Harris, 2008; Gou, Choudhury, & Benasich, 2011). In one study, EEG was recorded from healthy, typically developing 16-, 24-, and 36-month-old infants in the absence of a task (Benasich et al., 2008). Of particular interest in this study was oscillatory power in the gamma band, which refers to a measure of the strength of the signal in the gamma frequency range (~30–80 Hz). Activity in the gamma band has been found to be increased in a number of cognitive tasks and has been suggested to reflect neural synchronization critical to cognitive processing (e.g. Debener, Herrmann, Kranczioch, Gembris, & Engel, 2003; Herrmann & Knight, 2001; Herrmann, Munk, & Engel, 2004; Singer, 1999). Benasich et al. (2008) found that gamma power at frontal electrode sites distinguished between infants with a family history of language impairment and infants without such a family history, despite no language impairments being evident in the infant participants themselves at testing. Specifically, frontal gamma power was reduced in infants at risk for a language disorder compared with infants not at risk, and this difference became more pronounced over time (Benasich et al., 2008). In addition, correlational analyses across all infants showed that relatively high frontal gamma power during the EEG recording session was associated with relatively good performance on language production and comprehension tasks. These results highlight variability in the development of high-frequency oscillatory activity in the gamma band and its importance for language development (Benasich et al., 2008). In addition, this study provides a good example of how time-frequency approaches provide unique insight into our understanding of individual differences in language processing, and in this case, development, with potential clinical relevance.

Another line of research has focused on individual differences in oscillatory activity in the alpha band during semantic memory retrieval and language processing (~8–12 Hz) (Bornkessel, Fiebach, Friederici, & Schlesewsky, 2004; Hanslmayr, Sauseng, Doppelmayr, Schabus, & Klimesch, 2005; Klimesch, Schimke, & Pfurtscheller, 1993). In one study, Klimesch et al. (1993) recorded EEG while participants first viewed a string of letters and numbers (a memory set) and then were tasked with determining whether a subsequently presented target letter or number had appeared in the memory set. Of particular interest in this study were individual differences in the dominant frequency within the alpha band (e.g. alpha centered at 9 Hz vs. 11 Hz); this is termed individual alpha frequency (IAF). Individuals who performed relatively well on this semantic memory task were found to have a relatively high IAF (~10–12 Hz) compared with the IAF of participants who performed relatively poorly (~8–10 Hz) during the memory retrieval portion of the task (Klimesch et al., 1993).

This finding is particularly interesting in light of Klimesch's (2012) hypothesis outlining how shifts within a frequency band might serve to maximize (or minimize) cross-frequency coupling. Specifically, this hypothesis posits that due to the mathematical relationship between the centers of the traditional frequency bands (i.e. 5 Hz for theta and 10 Hz for alpha, which are harmonic frequencies), neural activity centered at 10 Hz would promote cross-frequency communication with theta-band activity, whereas a shift towards $<$10 Hz alpha activity would not promote (and might even prevent) this coupling (Klimesch, 2012). Since oscillatory activity in the theta frequency band has been strongly linked to memory processes (Gevins & Smith, 2000; Gevins, Smith, McEvoy, & Yu, 1997; Jensen, Gelfand, Kounios, & Lisman, 2002; Sauseng, Griesmayr, Freunberger, & Klimesch, 2010), it is therefore interesting that lower IAF has been associated with poor memory performance (Klimesch et al., 1993). In addition, Bornkessel et al. (2004) have shown that individuals with a relatively low IAF exhibit relatively large ERP effects of syntactic ambiguity during sentence reading compared with individuals with a relatively high IAF. Specifically, larger P600 effects, in which a greater positivity was found in response to syntactically ambiguous compared with syntactically unambiguous sentences, were found for high compared with low IAF readers (Bornkessel et al., 2004). Taken together, these results provide unique information about how changes in neural activity might relate to individual differences in language processing and comprehension, which would not be possible to obtain through other research methods.

As mentioned above, there are few studies to date that have used time-frequency approaches to investigate individual differences in language processing. However, this is a promising avenue for future research. One final example of how EEG can be used to contribute to this literature comes from a recent study in which EEG measures of oscillatory activity in the alpha frequency band were used to track attentional engagement during discourse (Boudewyn et al., in press). In this study, EEG was recorded while participants listened to short (4-sentence) stories in which two entities were introduced that were either very similar (e.g. two oaks), or less similar (e.g. one oak and one elm). This manipulation made a later anaphoric reference in a subsequent sentence (oak) ambiguous or unambiguous. ERPs measured at the anaphor yielded an Nref effect, an ERP effect that is sensitive to referential ambiguity of this type (Nieuwland, Otten, & Van Berkum, 2007; Nieuwland & Van Berkum, 2006, 2008; Van Berkum, 2008; Van Berkum, Brown, & Hagoort, 1999; Van Berkum, Brown, Hagoort, & Zwitserlood, 2003; see Van Berkum, Koornneef, Otten, & Nieuwland, 2007 for a review). Importantly, Nref amplitude at the anaphor was predicted by alpha power during the previous sentence in which its ambiguous or unambiguous antecedents had been established. Specifically, individuals with increased alpha power in ambiguous compared with unambiguous stories were less sensitive later on in the story to the ambiguity of the critical anaphor. Verbal ability as measured by a standardized vocabulary test was also predictive of greater sensitivity to referential ambiguity. These results suggest that relative increases in alpha power during listening comprehension were associated with attentional disengagement to the story, with processing consequences as the story continued (Boudewyn et al., in press). In addition, these results demonstrate how fluctuations in EEG power within- and between-participants can be used as a tool to track attentional engagement or other cognitive processes during language comprehension.

Section 3. ERPs/EEG and Individual Differences in Language Processing: Methodological Considerations

As noted in the Section on Individual Differences in Language Processing: A Short Background, language processing depends on a combination of specialized cognitive processes, domain-general processes, and experience. It is important to keep in mind, however, that these factors are all interrelated. In fact, many of the challenging aspects of designing or interpreting research on individual differences in language processing stem from this issue, which can be summarized very briefly: sources of variability are not independent. This means that it is problematic to conclude that one specific cognitive measure is driving observed variability in a language processing experiment if it is the only measure that was investigated.

This can be taken into account by a design that allows for the assessment of the impact of multiple sources of variability on the language process of interest. This has become the gold standard approach in the behavioral literature in this area: to collect multiple individual difference measures covering a range of known sources of variability in language comprehension, and to adopt analysis approaches that enable these to be simultaneously taken into account, such as multiple regression, hierarchical linear modeling, and structural equation modeling (e.g. Braze et al., 2007; Hamilton et al., 2013; Long, Prat, Johns, Morris, & Jonathan, 2008; McVay & Kane, 2012).

However, this approach has been utilized in very few ERP studies in this area (see Boudewyn et al., in press; Boudewyn, Long, & Swaab, 2012; 2013 for examples of some recent exceptions). Traditionally, ERP studies average across trials within a condition in order to achieve an acceptable signal-to-noise ratio, and then adopt a group comparison approach to investigate individual differences. A median split is often used, in which participants are divided into two groups based on the median score on some task (e.g. working memory span), and ERP effects in the two groups are compared. While this can be effective, it has two significant drawbacks. First, this approach turns a continuous variable (e.g. working memory span scores) into a categorical one (e.g. High vs. Low Span), thereby losing statistical power (Aiken, West, & Reno, 1991). Second, this approach does not allow for multiple sources of variance to be assessed when comparing the groups. For these reasons, it is important to use caution when interpreting data in which individual differences on a language processing task are based on a single measure, such as working memory span.[2]

This may seem like an understatement, given the issues highlighted above. However, it should be stressed that useful information can be (and has been) learned from such designs, and it is not always possible to collect data from an ideally wide range of measures from a large enough sample of participants to apply more sophisticated statistical techniques. Thus, it is recommended to use caution by applying appropriate limitations to the inferences that can be drawn from the results of a given study, rather than disregarding the results because a less-than-perfect design was used. That said, future ERP/EEG research will need to keep pace with the methodological and statistical developments being made in the behavioral literature in order to continue to provide novel insights into the neural processes underlying individual differences in language processing and comprehension.

Conclusions

ERP studies of individual differences in language processing are sometimes criticized as providing little more than conceptual replications of findings that have already been shown using behavioral measures. This is a fair criticism of some studies and is certainly something to keep in mind when designing future experiments. However, there are a number of electrophysiological studies that have gone beyond conceptual replications of behavioral work to add to our understanding of individual differences in how the brain processes language; a major goal of this review was to highlight some examples of these. The studies discussed above have fruitfully leveraged language-linked ERP effects (N400, P600, Nref) and other electrophysiological indices (slow waves, oscillations in the theta, alpha, and gamma frequency bands) to examine variability in language processing and in doing so have contributed unique insights to the field. This field continues to grow, with recent studies capitalizing on past behavioral work to make use of more sophisticated statistical modeling techniques. This will enable future research to be particularly informative in illuminating how language processing depends on a complex combination of specialized cognitive functions, domain-general processes, and experience.

Acknowledgement

Thank you to Erin Freed, Debra Long, and two anonymous reviewers for their thoughtful and very helpful comments on this manuscript.

Short Biography

Megan Boudewyn is a post-doctoral scholar at the University of California, Davis. Her research is aimed at understanding: (1) how context representations are used to aid in the processing of incoming words and phrases during real-time discourse comprehension, (2) the role of cognitive control and attention processes in contributing to individual differences in language comprehension; and (3) impairments in language processing and cognitive control in disorders such as schizophrenia.

Notes

* Correspondence address: Megan Boudewyn, Department of Psychology, University of California, Davis. E-mail: maboudewyn@ucdavis.edu

[1] There are many variations of WM span tasks. Perhaps the most common is the reading span task, originally developed by Daneman and Carpenter (1980). Participants in the reading span task read sets of sentences, generally ranging from a set size of two sentences through a set size of six sentences, and are asked to recall the sentence-final word from each sentence in a set. WM span in this case is the number of words correctly recalled; it is also sometimes calculated as the maximum set size for which a participant was able to correctly recall all sentence-final words. In other variations, participants are asked to complete a comprehension question after each sentence in a set, or solve a math equation.

[2] Aside from the broader issue of correlations with other individual differences measures, measures of working memory span also have some well-documented issues with reliability (Waters & Caplan, 1996, 2003). These are not the main focus of this section, but should also be taken into account; most typically, this is accomplished by including multiple working memory tasks and using a composite working memory score to increase reliability (Waters & Caplan, 2003).

Works Cited

Aiken, L. S., S. G. West, and Reno, R. R. (1991). *Multiple regression: testing and interpreting interactions: sage.*

Bell, L. C., and C. A Perfetti. (1994). Reading skill: some adult comparisons. *Journal of Educational Psychology,* 86. (2), 244–255. DOI: 10.1037/0022-0663.86.2.244

Benasich, A. A., Z. Gou, N. Choudhury, and K. D. Harris. (2008). Early cognitive and language skills are linked to resting frontal gamma power across the first 3 years. *Behavioural Brain Research,* 195. (2), 215–222.

Bornkessel, I. D., C. J. Fiebach, A. D. Friederici, and M. Schlesewsky. (2004). "Capacity" reconsidered: interindividual differences in language comprehension and individual alpha frequency. *Experimental Psychology,* 51. (4), 279–289.

Boudewyn, M. A., D. Long, M. J. Traxler, T. A. Lesh, S. Dave, G. R. Mangun, and T. Y. Swaab. (2015). EEG correlates of fluctuations in attention to discourse context predict sensitivity to referential ambiguity during language comprehension. *Journal of Cognitive Neuroscience.*

Boudewyn, M. A., D. L. Long, and T. Y. Swaab. (2012). *Cognitive control influences the use of meaning relations during spoken sentence comprehension.* Neuropsychologia.

——. (2013). Effects of working memory span on processing of lexical associations and congruence in spoken discourse. *Frontiers in Psychology,* 4.

Braze, D., W. Tabor, D. P. Shankweiler, and W. E. Mencl. (2007). Speaking up for vocabulary reading skill differences in young adults. *Journal of Learning Disabilities,* 40. (3), 226–243.

Clifton, C., M. J. Traxler, M. Taha Mohamed, R. S. Williams, R. K. Morris, and K. Rayner. (2003). The use of thematic role information in parsing: syntactic processing autonomy revisited. *Journal of Memory and Language,* 49.(3), 317–334.

Daneman, M., and P. A. Carpenter. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior,* 19. (4), 450–466.

Debener, S., C. S. Herrmann, C. Kranczioch, D. Gembris, and A. K. Engel. (2003). Top-down attentional processing enhances auditory evoked gamma band activity. *Neuroreport,* 14. (5), 683–686.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science,* 11. (1), 19–23.

Ericsson, K. A., and W. Kintsch. (1995). Long-term working memory. *Psychological Review,* 102. (2), 211.

Fiebach, C. J., M. Schlesewsky, and A. D. Friederici. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions. *Journal of Memory and Language,* 47. (2), 250–272.

Gernsbacher, M. A. (1993). Less skilled readers have less efficient suppression mechanisms. *Psychological Science,* 4(5), 294–298.

—— (1996). The structure-building framework: what it is, what it might also be, and why. *Models of understanding text,* 289–311.

—— (1997). Two decades of structure building. *Discourse processes,* 23. (3), 265–304.

Gernsbacher, M. A., and M. E. Faust. (1991). The mechanism of suppression: a component of general comprehension skill. *Journal of Experimental Psychology. Learning, Memory, and Cognition,* 17. (2), 245.

Gernsbacher, M. A., K. R. Varner, and M. E. Faust. (1990). Investigating differences in general comprehension skill*: Journal of Experimental Psychology. Learning, Memory, and Cognition,* 16. (3), 430.

Gevins, A., and M. E. Smith. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex,* 10. (9), 829–839.

Gevins, A., M. E. Smith, L. McEvoy, and D. Yu. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex,* 7. (4), 374–385.

Gou, Z., N. Choudhury, and A. A. Benasich. (2011). Resting frontal gamma power at 16, 24 and 36 months predicts individual differences in language and cognition at 4 and 5 years. *Behavioural Brain Research,* 220. (2), 263–270.

Hamilton, S. T., E. M. Freed, and D. L. Long. (2013). Modeling reader and text interactions during narrative comprehension: a test of the lexical quality hypothesis. *Discourse processes,* 50. (2), 139–163.

Hanslmayr, S., P. Sauseng, M. Doppelmayr, M. Schabus, and W. Klimesch. (2005). Increasing individual upper alpha power by neurofeedback improves cognitive performance in human subjects. *Applied Psychophysiology and Biofeedback,* 30. (1), 1–10.

Herrmann, C. S., and R. T. Knight. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neuroscience & Biobehavioral Reviews,* 25. (6), 465–476.

Herrmann, C. S., M. H. Munk, and A. K. Engel. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences,* 8. (8), 347–355. DOI:10.1016/j.tics.2004.06.006

Jensen, O., J. Gelfand, J. Kounios, and J. E. Lisman. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex,* 12. (8), 877–882.

Just, M. A., and P. A. Carpenter. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review,* 99. 122–149.

Kaan, E., A. Harris, E. Gibson, and P. Holcomb. (2000). The P600 as an index of syntactic integration difficulty. *Language and cognitive processes,* 15. (2), 159–201.

Kane, M. J., M. K. Bleckley, A. R. Conway, and R. W. Engle. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General,* 130. (2), 169.

Kane, M. J., and R. W. Engle. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General,* 132. (1), 47–70. DOI:10.1037/0096-3445.132.1.47

King, J. W., and M. A. Just. (1991). Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language,* 30. (5), 580–602. DOI:/10.1016/0749-596X(91)90027-H

King, J. W., and M. Kutas. (1995). Who did what and when? Using word-and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience,* 7. (3), 376–395.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review,* 95(2), 163.

Kintsch, W., and T. A. Van Dijk. (1978). Toward a model of text comprehension and production. *Psychological Review,* 85. (5), 363.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences,* 16. (12), 606–617. DOI:10.1016/j.tics.2012.10.007

Klimesch, W., H. Schimke, and G. Pfurtscheller. (1993). Alpha frequency, cognitive load and memory performance. *Brain Topography,* 5. (3), 241–251.

Kolk, H., and D. Chwilla. (2007). Late positivities in unusual situations. *Brain and Language,* 100. (3), 257–261. DOI:10.1016/j.bandl.2006.07.006

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

—— (2013). The proactive comprehender: what event-related potentials tell us about the dynamics of reading comprehension. In B. Miller, L. Cutting, P. McCardle (Eds.), Unraveling the behavioral, neurobiological, and genetic components of reading comprehension (pp. 176–192.). Baltimore: Paul Brookes Publishing.

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

——. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology,* 62. 621–647. DOI:10.1146/annurev.psych.093008.131123

Kutas, M., and C. Van Petten. (1994). Psycholinguistics electrified. *Handbook of psycholinguistics*, 83–143.

Kutas, M., C. Van Petten, R. Kluender. (2006). Psycholinguistics electrified II (1994–2005). *Handbook of psycholinguistics*, ed. by M. Traxler and M. A. Gernsbacher, 659–724: London, UK: Elsevier.

Long, D. L., C. Prat, C. Johns, P. Morris, and E. Jonathan. (2008). The importance of knowledge in vivid text memory: an individual-differences investigation of recollection and familiarity. *Psychonomic Bulletin & Review*, 15,(3), 604–609. DOI:10.3758/pbr.15.3.604

McVay, J. C., and M. J. Kane. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, 141. (2), 302.

Misyak, J. B., and M. H. Christiansen. (2012). Statistical learning and language: an individual differences study. *Language Learning*, 62. (1), 302–331.

Müller, H. M., J. W. King, and M. Kutas. (1997). Event-related potentials elicited by spoken relative clauses. *Cognitive Brain Research,* 5. (3), 193–203.

Münte, T. F., K. Schiltz, and M. Kutas. (1998). When temporal terms belie conceptual order. *Nature*, 395. (6697), 71–73.

Nakano, H., C. Saron, T. Y. Swaab. (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.

Nieuwland, M. S., M. Otten, and J. J. Van Berkum. (2007). Who are you talking about? Tracking discourse-level referential processing with event-related brain potentials. *Journal of Cognitive Neuroscience,* 19. (2), 228–236.

Nieuwland, M. S., and J. J. A. Van Berkum (2006). Individual differences and contextual bias in pronoun resolution: evidence from ERPs. *Brain Research*, 1118. (1), 155–167.

——. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: evidence from ERPs. *Brain and Language*, 106. (2), 119–131. DOI:10.1016/j.bandl.2008.05.001

Osterhout, L., A. Poliakov, K. Inoue, J. McLaughlin, G. Valentine, I. Pitkanen, and J. Hirschensohn. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics,* 21. (6), 509–521.

Perfetti, C. A. (2007). Reading ability: lexical quality to comprehension. *Scientific Studies of Reading,* 11. (4), 357–383. DOI:10.1080/10888430701530730

Perfetti, C. A., and L. Hart. (2002). The lexical quality hypothesis. *Precursors of functional literacy*, 11. 189–214.

Perfetti, C. A., E. W. Wlotko, and L. A. Hart. (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.

Sauseng, P., B. Griesmayr, R. Freunberger, and W. Klimesch. (2010). Control mechanisms in working memory: a possible function of EEG theta oscillations. *Neuroscience & Biobehavioral Reviews*, 34. (7), 1015–1022.

Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron*, 24. (1), 49–65.

Swaab, T. Y., K. Ledoux, C. C. Camblin, and M. A. Boudewyn. (2011). *Language-related ERP components Oxford Handbook of Event-Related Potential Components*. New York: Oxford University Press.

Tanner, D., K. Inoue, and L. Osterhout. (2014). Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, 17. (02), 277–293.

Tanner, D., J. McLaughlin, J. Herschensohn, and L. Osterhout. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, 16. (02), 367–382.

Tanner, D., L. Osterhout, and J. Herschensohn (2009). Snapshots of grammaticalization: differential electrophysiological responses to grammatical anomalies with increasing L2 exposure. Paper presented at the Proceedings of the 33rd Boston University conference on language development.

Traxler, M. J., R. S. Williams, S. A. Blozis, and R. K. Morris. (2005). Working memory, animacy, and verb class in the processing of relative clauses . *Journal of Memory and Language,* 53. (2), 204–224. DOI:10.1016/j.jml.2005.02.010

Van Berkum, J. J. A. (2008). Understanding sentences in context what brain waves can tell us. *Current Directions in Psychological Science,* 17. (6), 376–380.

Van Berkum, J. J. A., C. M. Brown, and P Hagoort. (1999). Early referential context effects in sentence processing: evidence from event-related brain potentials. *Journal of Memory and Language,* 41(2), 147–182.

Van Berkum, J. J. A., C. M. Brown, P. Hagoort, and P. Zwitserlood. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology*, 40. (2), 235–248.

Van Berkum, J. J. A., A. W. Koornneef, M. Otten, and M. S. Nieuwland. (2007). Establishing reference in language comprehension: an electrophysiological perspective. *Brain Research,* 1146. 158–171. DOI:10.1016/j.brainres.2006.06.091

Van Dijk, T. A., and W. Kintsch. (1983). *Strategies of discourse comprehension*: Academic Press New York.

Van Dyke, J. A., and C. L. Johns. (2012). Memory interference as a determinant of language comprehension. *Language and linguistics compass,* 6. (4), 193–211.

Van Dyke, J. A., C. L. Johns, and A. Kukona. (2014). Low working memory capacity is only spuriously related to poor reading comprehension. *Cognition,* 131. (3), 373–403. DOI:10.1016/j.cognition.2014.01.007

Vos, S. H., and A. D. Friederici. (2003). Intersentential syntactic context effects on comprehension: the role of working memory. *Cognitive Brain Research*, 16. (1), 111–122.

Vos, S. H., T. C. Gunter, H. H. Kolk, and G. Mulder. (2001). Working memory constraints on syntactic processing: an electrophysiological investigation. *Psychophysiology*, 38.(01), 41–63.

Waters, G. S., and D. Caplan. (1996). *The capacity theory of sentence comprehension: critique of Just and Carpenter (1992).*

——. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*, 35. (4), 550–564.