

The neural mechanisms of coreference

Kerry Ledoux^{1,2} and C. Christine Camblin³

¹Cognitive Neurology/Neuropsychology, ²Department of Neurology,
The Johns Hopkins University School of Medicine, Baltimore, MD, USA

³Department of Social and Behavioral Sciences,
California State University, Monterey Bay, CA, USA

Corresponding Author:
Kerry Ledoux
Cognitive Neurology/Neuropsychology
Johns Hopkins University
1629 Thames St. Suite 350
Baltimore, MD 21231

Email: kledoux1@jhmi.edu
Phone: (443) 287-1712
Fax: (410) 955-0188

The neural mechanisms of coreference

Coreference (the mechanism by which two linguistic expressions are taken to refer to the same entity in the world) is a universal and essential feature of discourse. Without this tool, our ability to comprehend language would be severely impaired. Due to their central role in discourse comprehension, the processes by which coreference is established have been the focus of numerous behavioral studies. In this article, we review studies that build upon that body of work by employing the event-related potential technique to elucidate the neuronal bases underlying the representation and processing of discourse. We include in our review studies which violate the formal constraints on the establish of coreference described by linguistic theory; studies that examine the relative ease or difficulty of establishing coreference under different conditions as predicted by processing models; and studies that examine the modulation of lexico-semantic processing (such as priming) by processes associated with the establishment of coreference. Additionally, we discuss the implications of this research for models of coreference.

In nearly every aspect of our lives, our daily behavior depends upon the comprehension of language. The goal in language processing is to develop a coherent, cohesive interpretation of the linguistic input. Although it is convenient (at least in the experimental realm) to think of language processing in terms of isolated words or sentences, natural language processing rarely occurs at either of these levels alone. Instead, the words or sentences that one reads or hears are embedded in a larger context, the discourse, the distinct properties of which have unique and direct processing consequences. This paper is concerned with a specific set of discourse processes, those engaged in the service of the establishment of coreference, and with the neural consequences of such processing.

Two linguistic expressions are said to be *coreferential* if they refer to the same entity in the world; the first expression (the *antecedent*) introduces that entity into the discourse model, and the second expression (the *anaphor*) refers to it. The process of establishing coreference between an anaphor and its antecedent is essential to constructing a message-level representation of a discourse. Take, for example, the following passage:

Over the summer, Janet drove her son to the lake so that he could go swimming. She dropped him off at noon, and he spent the rest of the day with his friends. When she picked him up, Evan told her all about the games they had played in the water. Janet is thrilled that Evan has made so many friends this summer.

The comprehension of this passage depends on the establishment of coreferential relations. One must determine that the anaphor *he* in the first sentence refers to the same entity in the world as the antecedent expression *her son*, and that *they* in the third sentence refers back to *friends* in the

second sentence. As one can also see from the example, pronouns are not the only means by which coreference can be established; at one point, the proper name *Evan* is used to refer back to the noun phrase *her son*, and at another, the reader must determine that the two instances of *Evan* refer to the same person. Thus, we see that coreference can occur both within and between sentences and can be established by the use of full expressions (such as descriptions or names) as well as reduced expressions (such as pronouns and ellipses).

Scores of behavioral studies have suggested that the establishment of coreference involves the coordination of processes acting on the antecedent and on the anaphoric expression. That is, processing coreferential relationships between different expressions first requires the establishment and maintenance of a representation of the information conveyed by the antecedent expression. This stored information must then be retrieved based on cues in the anaphoric expression. These comprehension processes of representation and retrieval, as well as the production process of choosing an appropriate coreferential expression, are strongly influenced by syntactic structure, discourse principles, and pragmatic constraints. The results of behavioral studies have led to the development of multiple prominent theories of coreferential processing; while we will briefly discuss a handful of these in our review, the interested reader is pointed toward several excellent models of this kind (among them Ariel 1990; Gernsbacher 1990; Greene, McKoon, & Ratcliff 1992; Gordon & Hendrick 1998; Sanford & Garrod 1998; Almor 1999; Garnham 2001; Badecker & Straub 2002).

The use of techniques that assess neural activity to inform questions of cognitive function has offered a new perspective on many issues in language processing. Two techniques in particular – event-related brain potentials (ERPs) and functional neuroimaging – have contributed to our understanding of the cognitive and neural architectures of language. The two

methods are often described as offering nearly complementary benefits: ERPs provide information about neural processing with very fine temporal resolution, but relatively poor spatial resolution. Functional neuroimaging, on the other hand, offers relatively fine spatial resolution, but, because of the sluggish relationship between neural activity and the hemodynamic response, it provides little information about the timing of rapid cognitive events. To date, few functional neuroimaging studies have examined processes of coreference, and while those that have (for example, Hammer, Goebel, Schwarzbach, Münte, & Jansma 2007; Nieuwland, Petersson, & Van Berkum 2007) have contributed to our understanding of the localization of such processing and have laid the groundwork for future research, the small number of studies makes generalizations about coreference difficult. For that reason, we have chosen in this review to focus on the relatively larger number of studies of coreference that have used the ERP technique. Below, we briefly describe this technique, and some of its basic findings regarding the neural representation and processing of language.

Event-related brain potentials (ERPs) are measures of neural activity that are time-locked to some external, triggering event. ERPs are derived from recordings of electrical activity made at the scalp (from the electroencephalogram, or EEG). ERPs can be obtained for spoken and written language, without the necessity of a behavioral task that is extraneous to the purposes of natural language comprehension. They also provide a continuous neural measure that can tap into language processes as they unfold in real time. Because separable ERP components are sensitive to separable aspects of language processing (e.g., semantics vs. syntax), they can give a direct indication of the component mechanisms of language comprehension processes.

Several ERP components have been associated with different aspects of language processing. The N400 is a negative deflection in the ERP waveform that peaks approximately

400 ms post-stimulus onset, and is maximal over posterior electrode sites. A reduction of the amplitude of the N400 is found to words that can be easily integrated into the preceding word, sentence or discourse context, and as such, this component is considered sensitive to processes of semantic integration (e.g., Kutas & Hillyard 1980; Brown & Hagoort 1993; Holcomb 1993).

The P600 is a positive deflection in the ERP waveform that begins approximately 500 ms after the onset of a word, and which can last several hundred milliseconds. The P600 has traditionally been associated with processes of syntactic structure building or with processes of syntactic analysis or reanalysis (Osterhout & Holcomb 1992; Hagoort, Brown, & Groothusen 1993; Friederici 1998). However, recently, several groups (Kim & Osterhout 2005; van Herten, Kolk, & Chwilla 2005; Kemmerer, Weber-Fox, Price, Zdanczyk, & Way 2006; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb 2006) have observed P600-like effects to seemingly semantic violations (*For breakfast the eggs would eat only toast and jam*), casting some doubt on the exact functional significance of the P600 component. It may be that the P600 reflects an inconsistency between semantic information and combinatorial information (Kuperberg 2007).

Alternatively, this component may reflect the engagement of executive control processes (van Herten, Chwilla, & Kolk 2006; Kolk & Chwilla 2007). Finally, a negative shift in the ERP waveform at (primarily left) anterior electrode sites during approximately the same time window as the N400 has been labeled the left anterior negativity (LAN). The functional significance of the LAN is a matter of some debate. Some have associated it with processes of working memory, including the working memory that is recruited in the service of language processing (Kluender & Kutas, 1993; King & Kutas 1995). Others have described the LAN as an index of some aspect of difficulty during syntactic processing (Friederici 1995, 2002; Hagoort 2003).

In the following sections, we provide a brief review of studies that have used event-related potentials to study coreference, and attempt to summarize the insights that have been gained from such studies. The nature of the ERP technique is such that differences in cognitive processing are reflected in *relative* changes in the signal of interest in different experimental conditions. So, for example, a researcher might report an N400 ERP effect even if the absolute measurement of the EEG signal is positive, as long as the waveform in one condition is more or less negative than that in another. Thus, the functional interpretation of ERP data requires comparisons between conditions that differ along functional cognitive dimensions of interest. We divide our review into three major sections, depending on the types of stimuli that were used in the conditions of comparison. In the first section, we review studies in which formal constraints on the establishment of coreference (as described by linguistic or psycholinguistic models) are violated. In the second section, we review studies in which theoretical models have been used to motivate predictions about the relative ease or difficulty of establishing coreference in stimuli that do not contain such formal violations. In the third section, we review studies that have examined the ways in which lexical processing is altered by the additional demands of the processing of coreference. We provide brief attempts to relate the results of these ERP studies to the relevant theoretical and behavioral linguistic or psycholinguistic work; for deeper discussion of these issues, the interested reader is encouraged to consult the original sources.

Violations of Formal Constraints on the Establishment of Coreference

Agreement violations. According to linguistic theory, there are several minimal constraints that govern the establishment of coreference between an anaphor and its antecedent. At a basic level, an anaphor must match its antecedent on certain elemental features, such as

number, animacy, and (conceptual/biological and syntactic, when available) gender. In fact, most linguistic models suggest that coreference cannot be established if there is no available potential antecedent that matches the anaphor on these elemental features, a claim that is supported by much behavioral evidence (Ehrlich 1980; Vonk, 1984; Garnham & Oakhill 1985; Vonk, 1985; McDonald & MacWhinney 1995)

In general, violations of agreement between an anaphor and its antecedent (like other types of agreement violation) result in a larger amplitude P600 when compared to well-formed control stimuli. This P600 effect is sometimes accompanied by a LAN effect (with a larger LAN reported to violations than to controls), and may be accompanied by differential ERP effects at words downstream from the point of coreference. One of the earliest demonstrations of this came from Osterhout and Mobley (1995; see also Osterhout 1997), who reported P600 effects to morphosyntactic agreement violations. Participants read sentences with or without violations of reflexive-antecedent number agreement (*The hungry guests helped themselves/*himself to the food*) or reflexive-antecedent gender agreement (*The successful woman congratulated herself/*himself on the promotion*). A larger P600 was observed to the reflexive pronoun in both violation conditions, relative to the controls. Additionally, a broad, N400-like negativity was observed to the final word of sentences that contained a violation, relative to sentences that did not, suggesting that readers may have experienced difficulty with semantic integration in the presence of a morphosyntactic violation. In a subsequent study, Osterhout, Bersick, and McLaughlin (1997) contrasted these definitional gender agreement violations with stereotypical gender agreement violations, such as *The doctor prepared himself/herself for the operation*. In these cases, the reflexive pronoun disagreed with the gender stereotypical of the descriptive noun antecedent. Both definitional violations and stereotypical violations resulted in an increased

P600 amplitude relative to the non-violating controls; this P600 effect was larger for the definitional than for the stereotype condition. As in the previous study, an N400 effect was found at the final word for sentences that contained a gender agreement violation, but only in the definitional condition. The authors suggested that this pattern of results may have arisen because in the stereotype condition, the reflexive pronouns are not violations of *formal* morphosyntactic principles, but instead of processing *preferences* that are informed by stereotypes. This may make syntactic reanalysis at the pronoun less demanding than in the definitional condition (resulting in a smaller P600), and may alleviate the integration difficulty experienced at the end of sentences that contained a true violation.

Occasionally, P600 effects to morphosyntactic agreement violations have been accompanied by a left anterior negativity. For example, Coulson, King, and Kutas (1998) manipulated the probability of the occurrence of ungrammatical sentences within a block of stimuli to further elucidate the functional significance of the P600 component of the ERP. Within this framework, they included a pronoun condition, in which well-formed sentences (*Ray fell down and skinned his knee*) were compared to those that contained a violation of overt case marking on pronouns (*Ray fell down and skinned *he knee*). Relative to grammatical pronouns, ungrammatical pronouns elicited a larger P600 and a larger LAN, the latter of which the authors argued reflects working memory operations involved in sentence processing.

Binding violations. According to linguistic theory, different constraints govern the establishment of coreference with different referring expressions. For example, Chomsky's Government and Binding Theory (1981) differentiates between the constraints on coreference with reflexives (like *himself* or *themselves*), pronouns (like *she* or *it*), and full referring expressions (like names or other noun phrases). Specifically, a reflexive is said to be bound in

its syntactic domain, meaning that its antecedent must occur in the immediate, local context (for instance, the same clause). Pronouns and full referring expressions, on the other hand, must not be bound in this way, meaning that their antecedents cannot occur this locally if coreference is to be established.

In the preceding agreement violation studies, the use of reflexives necessitated the location of an antecedent within the local discourse context, according to binding principles. Because no appropriate antecedent (one matching the reflexive on morphosyntactic features) was available in that local context, the reflexive could be considered a formal violation. Other studies have used pronoun anaphors, in sentences such as *John gave Bill some money so she could buy some candy*. Because pronouns are not bound to the immediate local context, a reader or listener, upon encountering *she* in such a sentence, could look outside of the local sentence or discourse context to find an appropriate antecedent for the pronoun. In this case, then, the use of *she* does not result in a formal violation of binding principles. It may, however, result in a *perceived* violation (if readers or listeners prefer to find an antecedent within the local context), in which case it may be taken to be an instance of referential failure nonetheless.

ERP studies of such pronouns have suggested that this is indeed the case. Osterhout and Mobley (1995), who examined reflexive (bound) pronoun processing in their Experiment 1, conducted a second experiment to see if similar ERP effects would be found for free pronouns. Participants read sentences like *The aunt heard that she/he had won the lottery*, in which pronouns agreed or disagreed in gender with the local, available antecedent. As with the reflexive pronouns in Experiment 1, there was a P600 effect to violating pronouns and an N400 effect at the sentence-final word of violation stimuli. The P600 effect in this case, however, was smaller in magnitude than that seen in the first experiment, leading the authors to wonder if

individual differences had played a role in Experiment 2. That is, perhaps participants differed in their engagement of processes to find an appropriate antecedent of the pronoun, with some participants being more likely than others to look outside the local sentence context, assuming the existence of another person related to the story. They grouped the participants' ERPs based on their judgments in the behavioral task, with one group labeling these sentences as generally unacceptable, and another labeling them as generally acceptable. For participants who judged the sentences to be unacceptable, there was a large P600 effect at the pronoun; for participants who judged the sentences to be acceptable, no such effect was observed, supporting the hypothesis of individual differences. Subsequent studies have replicated the P600 effect to referentially failing pronouns, using both auditory (Van Berkum, Zwisterlood, Bastiaansen, Brown & Hagoort 2004) and visual (Nieuwland & Van Berkum 2006) presentation of sentences such as *Anna shot at Linda as he jumped over the fence*.

In English, gender marking is reserved for pronouns that refer back to entities that are biologically male or female. In many languages, though, marking is determined not only by biological gender (for those entities that exhibit it), but also by syntactic gender (determined arbitrarily within and across languages). In French, for example, *bras* (*arm*) is masculine, but *jambe* (*leg*) is feminine. The syntactic gender of a word will often be marked on any other words in the sentence that are related to that word: determiners, modifying adjectives, and pronouns. A series of studies (Schmitt, Lamers, & Münte 2002; Hammer, Jansma, Lamers, & Münte 2005) has taken advantage of this linguistic feature to examine the relative contribution of semantic and syntactic information during the processing of pronoun anaphors that did or did not violate syntactic or conceptual agreement with their antecedents. Schmitt and colleagues (2002) examined the processing of German non-diminutives (*Bub/boy*) and diminutives (*Bübchen/little*

boy). For non-diminutives, conceptual/biological and syntactic gender match: for example, *Bub* both refers to an entity with male gender semantically, and also carries a masculine syntactic gender marking. For diminutives, on the other hand, that is not the case; for *Bübchen*, the conceptual/semantic feature (male) differs from the syntactic feature (neuter). Sentences were developed in which a pronoun referred to either a diminutive or a non-diminutive antecedent. The pronoun was manipulated to match the antecedent on semantic/biological gender only; on syntactic gender only; on both semantic and syntactic gender; or on neither. All gender violations (whether conceptual/semantic or syntactic) resulted in larger P600s (relative to the condition in which both types of gender feature matched the antecedent). For non-diminutives only, double violations (violations of both semantic and syntactic gender) were also accompanied by N400 effects; this was not seen for diminutives, either for double violations or for single violations, suggesting that the establishment of coreference with diminutive antecedents did not depend differentially on semantic information.

In a subsequent study (Hammer, et al. 2005), the authors continued this line of inquiry by examining the processing of coreference with different types of antecedents in German. In this case, they developed sentences in which the antecedent was either a person or a thing. For person antecedents, an anomalous pronoun mismatched the antecedent on both conceptual gender and on syntactic gender. For thing antecedents, an anomalous pronoun mismatched the antecedent on syntactic gender only. A comparison of pronoun violations to controls revealed a P600 effect that was larger for person antecedents than for thing antecedents. Additionally, at the word following the pronoun, a larger amplitude N400 was observed for violation sentences, but only in the case of thing antecedents, suggesting that in this condition (but not in the case of

person antecedents), readers experienced additional integration difficulties immediately downstream from the pronoun.

Other studies have compared the processing of referring expressions that are said to differ in their binding constraints. One such study examined the processing of the same lexical forms in different structural binding contexts (Harris, Wexler, & Holcomb 2000). As described above, when reflexives are found in argument position of a verb (as in *Susan liked herself*), they must refer to an antecedent in the local domain, in exact opposition to a pronoun (as in *Susan liked her*), which *cannot* refer to the local antecedent. When reflexives are found in a non-argument position, they are called logophors, and they can, in this case, on grammatical grounds, be used interchangeably with pronouns (*Susan liked to take pictures of Michael and herself/her*). Pragmatic constraints, however, may influence the choice of logophor or pronoun in such sentences. Harris and colleagues (2000) conducted an ERP experiment to explore the processing of reflexives and logophors. Participants read reflexive sentences like *The pilot's mechanics brow-beat themselves/*himself after the race* (in which the singular reflexive was considered to be a formal violation, because it does not match the only available antecedent, *mechanics*, in number); and logophoric sentences like *The pilot's mechanics brow-beat Paxton and themselves/himself after the race* (in which both forms are grammatically allowed, but in which the plural form may be preferred for the pragmatic reason of agreeing with the subject of the sentence). The ERP results confirmed the authors' predictions about the nature of the perceived violations in the sentences: for reflexives, a comparison at the pronoun revealed a larger P600 for the formal violation, a difference that was not seen for the logophors. Instead, at the logophor, the singular form, relative to the plural form, elicited a marginally more negative response in the N400 time-window. The P600 difference for reflexives, then, suggests the importance of

syntactic information in their processing. The processing of logophors, on the other hand, is governed not by syntactic principles, but by principles of pragmatics.

The absence of an explicit antecedent. A recent study (Filik, Sanford, & Leuthold 2007) examined the processing of pronouns that did not have explicit antecedents. According to linguistic theory, such pronouns should generally result in referential failure, with an associated processing cost. However, in colloquial English, it is not uncommon to use (without communication cost) constructions like *At school today they made an announcement about the field trip*, in which no explicit antecedent is provided. Event-related potential measures were used to determine whether actual processing costs exist in such cases. Participants read sentences with singular (*he/she*) or plural pronouns (*they*), in the second sentence of two-sentence stimuli that either did or did not introduce an explicit antecedent in the first sentence. As expected, singular pronouns that were not preceded by an explicit antecedent engendered a processing cost, as evidenced by a larger P600 component. Such a cost was not elicited by plural pronouns without an explicit antecedent, confirming that indeed, the syntactic processor seems able to handle such technically anomalous pronouns without difficulty.

To summarize, violations of the constraints on the establishment of coreference (in the form of morphosyntactic agreement violations, binding relations violations, or the absence of an explicit antecedent) are detected very quickly, within approximately 500 ms of the presentation of the anaphor. These anomalous anaphors seem to be treated by the processor as violations of syntax, not semantics, as evidenced by the differential effects on the P600, a component that (at least traditionally) has been associated with syntactic (re)analysis or repair. This effect can be accompanied by a left anterior negativity, which may reflect difficulties in syntactic integration or the increased engagement of working memory resources during the processing of violations.

The difficulty evidenced by the P600 effect may be accompanied by an N400 effect, observed downstream from the anaphor, which suggests that ongoing processes of semantic integration can be hampered by a syntactic violation that disrupts the establishment of coreference. Finally, results suggest that in some cases, the magnitude of these ERP effects may be influenced by individual differences in language processing.

Factors Affecting the Establishment of Coreference in the Absence of a Formal Violation

In the previous section, we reviewed studies of factors that (according to linguistic models) formally prohibit coreference. In this section, we turn our attention to studies of situations in which the establishment of coreference should be possible, and indeed may be possible using multiple forms of referring expression. In these cases, experiments have been motivated by psycholinguistic models of processing in which factors that affect the ease or difficulty of establishing coreference using different referring expressions are described. Some of these factors relate to the organization of the discourse itself. Others relate to the availability of potential antecedents in the mental representation of a discourse and the differential effects this has on processing different forms of referring expressions. Finally, we include mention of one semantic factor (implicit causality) and its role in determining the ease of establishing coreference.

Referential ambiguity. One feature that can affect coreferential processing is the number of potential antecedents that match the anaphor that are present in the (local) discourse context. When several candidate antecedents match an anaphor on the essential elements described previously, and when binding or other constraints do not eliminate some of these candidates from consideration, an instance of referential ambiguity may arise. In a series of papers, Van

Berkum and colleagues (Van Berkum, Brown, & Hagoort 1999; Van Berkum, Brown, Hagoort, & Zwitserlood 2003; Van Berkum, et al. 2004; Nieuwland & Van Berkum 2006; Nieuwland, Otten, & Van Berkum 2007; for review, see Van Berkum, Koornneef, Otten, & Nieuwland 2007) have shown that situations of referential ambiguity (when compared with situations of referential success) consistently elicit an ERP response that is distinctly different from that observed in situations of referential failure. They have called this difference the Nref effect: a frontal negative shift that emerges rapidly (within 300 ms) after the presentation of the ambiguous (but not the unambiguous) anaphor, and that is frequently sustained over several hundred milliseconds. In the first demonstration of this effect (Van Berkum, et al. 1999), participants read three-sentence stories in which a main character interacted with two other characters, one of whom was subsequently referred to by a repeated noun phrase anaphor (for example, *the girl*). In the one-referent condition, the antecedent of this anaphor was unambiguous, in that it could only refer to one of the two characters who had been introduced (ex., *David had told the boy and the girl...*). In the two-referent condition, the antecedent of the anaphor was ambiguous (ex., *David had told the two girls...*). The ERP elicited by the repeated NP in the two-referent condition was more negative at frontal sites than that elicited by the repeated NP in the one-referent condition, a shift that onset approximately 280 ms after the presentation of the critical word. The authors speculated that such an effect might be specific to referential ambiguity; alternatively, it might be related to the LAN effect and as such, might reflect the greater working memory demands of maintaining information about and selecting between two potential referents as opposed to one. This finding was also shown for the same stimuli presented in the auditory modality (Van Berkum, et al. 2003). Further research (Van Berkum, et al. 2004) demonstrated a similar effect for referentially ambiguous pronouns in a single-sentence context

(*David shot at John/Linda as he jumped over the fence*). Subsequently, the magnitude of the Nref effect was shown to depend upon individual differences in language processing ability: readers who scored higher on a language working memory span task also tended to show larger negative shifts to referentially ambiguous pronouns than did readers with lower span scores (Nieuwland & Van Berkum, 2006). Importantly, this effect of referential ambiguity has been demonstrated to reflect deep situation model ambiguity, as opposed to superficial textbase ambiguity (Nieuwland, Otten, & Van Berkum 2007).

Parallelism. A parallel discourse structure is one in which the anaphor appears in the same grammatical role as the antecedent. According to the *parallel function strategy*, anaphors are expected to refer back to the entity in the previous sentence that shares its grammatical role. So, for example, an anaphor in subject position may be preferentially interpreted as coreferential with the antecedent found in subject position of the preceding clause or sentence (Sheldon 1974; Grober, Beardsley, & Caramazza 1978; Erhlich 1980; Smyth 1994; Gordon & Scearce 1995; Chambers & Smyth 1998). The lack of parallelism can hinder coreferential processing, an effect that was demonstrated in an ERP study in which pronouns and repeated names in subject and object position referred to antecedents in subject or object position in a preceding sentence. When compared to the ERPs to critical words in parallel structures, ERPs to pronouns and repeated names in non-parallel structures exhibited an increased negativity over parietal electrode sites, suggesting an N400 effect (Streb, Rösler, & Hennighausen 1999).

Distance. The establishment of coreference between an anaphor and its antecedent may also be easier when the two appear more closely together in a discourse. That is, coreference may be more readily established if the antecedent expression has been mentioned recently in the discourse, relative to the appearance of the anaphor (Garrod & Sanford 1977; Clark & Sengul

1979; Daneman & Carpenter 1980; Ehrlich & Rayner 1983). Streb, Hennighausen, and Rösler (2004) compared the ERP response to pronouns and coreferential repeated names at different linguistic distances from their antecedents. In the near condition, the antecedent of the anaphor was in the previous sentence; in the medium condition, it was two sentences back; and in the far condition, three sentences back. For each type of anaphor, ERPs in the medium and far conditions showed an increase in the amplitude of the N400 relative to the near condition, suggesting that the integration of both pronouns and repeated names became more difficult with distance from the antecedent, and to a similar extent for the two types of expressions.

Availability of the antecedent. Linguists and psycholinguists have observed differences in the distribution of different types of referring expressions in language. One source of these differences has been described generally as the availability of the antecedent within the mental representation of the discourse: pronouns tend to be used to refer to antecedents that are more available in the mental model, while full referring expressions (such as repeated names and noun phrases) tend to be used to corefer with less available antecedents. This characterization, however, begs the question of what makes an antecedent more or less available in the discourse representation.

Several processing models have attempted to describe features of the discourse that might determine the availability of the antecedent. These models have used various terms to refer to “availability” (including discourse prominence, salience, focus, topicality, accessibility, and activation), but what they have in common is their proposal of (often the interaction of multiple) factors that determine the availability of the antecedent. For example, Ariel’s Accessibility Hierarchy (1988, 1990) characterized antecedent availability along four dimensions: distance between the anaphor and the antecedent; competition (in terms of the presence of other potential

antecedents); saliency (whether the antecedent is considered the topic of the discourse or not); and unity (whether the antecedent is in the same segment or frame as the anaphor). Generally, according to this hierarchy, pronouns would tend to corefer with antecedents that were low on distance and competition but high on saliency and unity, and full referring expressions the opposite.

Other models place more emphasis on the structural characteristics of the sentence or discourse in determining the availability of antecedents in the discourse model. For example, according to Centering Theory (Gordon, Grosz, & Gilliom 1993; Grosz, Joshi, & Weinstein 1995), potential antecedents (called forward-looking centers) are ranked according to grammatical function; the highest ranked forward-looking center is the preferred center. Also among the potential antecedents is one known as the backward-looking center, as it refers back to an entity mentioned previously in the discourse. When the backward-looking center is the same as the preferred center, the topic of the discourse is continued, and this entity becomes the most available potential antecedent, to which a pronoun can readily corefer. In situations requiring a topic shift (those in which the backward-looking center is not the same as the preferred center), several antecedents may be equally available, and a fuller referring expression will be processed more easily.

Gordon and Hendrick's Discourse Prominence Theory (1998) builds on Centering Theory by describing different processing operations that act on pronouns and full referring expressions. Coreferentially repeated names generally require additional processing than pronouns, as the establishment of coreference between two names requires an additional construction rule (or processing operation) to establish the equivalence of two entities in the discourse model predicated on the same name. Also, as in Centering Theory, potential

antecedents are ordered in terms of “prominence” in the discourse representation (determined by factors such as grammatical function or the structural depth of embedding). The processing of pronouns and repeated names proceeds differently according to prominence: pronouns trigger a search through entities in the model in descending order of prominence, while repeated names trigger a search for an antecedent in ascending order of prominence.

Finally, according to Almor’s Informational Load Hypothesis (1999; Almor & Nair 2007), the establishment of coreference with different referring expressions involves a balance between discourse function (the identification of an antecedent and the addition of new information about it to the discourse) and processing cost. The activation of an anaphor in working memory and the reinstatement of the representation of the antecedent contribute to this processing cost. Repeated names and noun phrases also incur an additional processing cost that is associated with the potential they introduce for semantic interference between the identical working memory representations of the antecedent and the anaphor that become simultaneously active before they are integrated. The high processing cost associated with the activation of both a repeated expression and a highly accessible antecedent must be balanced by a specific discourse function (say, adding emphasis); otherwise, a processing cost will be incurred.

While behavioral studies have provided evidence for several of these availability models (see, for example, Gordon, et al. 1993; Almor 1999; Gordon, Hendrick, Ledoux, & Yang 1999), event-related potential studies to date have primarily addressed the predictions of Discourse Prominence Theory (DPT). Gordon and colleagues (Swaab, Camblin, & Gordon 2004; Camblin, Ledoux, Boudewyn, Gordon, & Swaab 2007; Ledoux, Gordon, Camblin, & Swaab 2007) have explicitly examined the differential influence of prominence for pronouns and for coreferential repeated names. Remember that within DPT, the ease of establishing coreference with anaphoric

pronouns is *increased* by antecedent prominence; conversely, the ease of establishing coreference with repeated names is *decreased* by antecedent prominence. This prediction was first tested in an ERP experiment that examined the processing of coreferential repeated names and pronouns (Swaab, et al. 2004). Participants read sentences in which coreference was established by either a pronoun or a repeated name in the second clause of the sentence. This anaphor referred to an antecedent that was either prominent in the discourse representation (i.e., the singular subject of the sentence, as in *Pam washed the dishes while Pam/she talked about politics*) or not (i.e., an entity that was more deeply structurally embedded as a member of a conjoined subject, as in *Pam and Joe washed the dishes while Pam/she talked about politics*). Compared to the ERP to the coreferential repeated name when the antecedent was prominent in the discourse representation, the ERP to the repeated name in the non-prominent condition showed a reduction of the N400, suggesting that the difficulty of establishing coreference with a repeated name increased with the prominence of the antecedent. The comparison of pronouns to repeated names directly, however, was made difficult by differences between the two types of referring expressions along several lexical dimensions (frequency, length, word class) that have been shown to have an influence on ERPs. A subsequent study (Ledoux, et al. 2007) compared the processing of coreferential repeated names to that of new names (which introduced a new entity into the discourse model) in the same sentences and at the same sentence position. Sentences introduced a singular (prominent) or conjoined (non-prominent) first noun phrase, followed by a coreferential repeated name (*Pam [and Joe] washed the dishes while Pam talked about politics*) or a new name (*Pam [and Joe] washed the dishes while Ann talked about politics*). For repeated names (but not for new names), the amplitude of the N400 ERP component was reduced in the non-prominent relative to the prominent condition, suggesting

that the processing of a coreferential repeated name was more difficult when the antecedent was prominent in the discourse representation. These results were later replicated with auditory presentation (Camblin, et al. 2007). These findings support a model of coreference such as the DPT by showing that the ease or difficulty of establishing coreference using repeated names varied depending on the accessibility of the antecedent, when accessibility was defined in terms of structural characteristics such as the depth of linguistic embedding.

Nair, Almor, and Vendemia (2007) extended this work by examining the differential neural response to pronominal and repeated name coreference within and between sentences. For the discourse (between sentence) condition, stimuli like those used by Gordon and colleagues above were split into two sentences, as in *John (and Mary) went to the store. John/He wished to buy some candy.* The larger N400 to repeated names that followed a prominent antecedent was replicated, but, interestingly, only for the within-sentence condition. A larger LAN was observed for the conjoined condition relative to the singular condition for repeated names (but not for pronouns), suggesting a greater involvement of working memory resources in repeated name coreference when the antecedent was part of a conjoined noun phrase. Finally, a larger LAN was observed in the between-sentence condition relative to the within-sentence condition, suggesting the greater recruitment of working memory resources when coreferential relations span more than one sentence.

Implicit causality. One semantic feature that has been demonstrated to influence the establishment of coreference is verb implicit causality. Implicit causality is a lexical-semantic feature of certain interpersonal verbs by which information about the cause of events described by a verb is conveyed implicitly as part of the verb's meaning. For example, when asked to complete a sentence fragment such as *Gregory apologized to Allison because...*, most people

will continue the sentence by saying something about Gregory, not about Allison. The verb *apologize (to)* is thus said to be biased towards its first noun phrase (NP1). The verb *thank on the other hand* (as in *Gregory thanked Allison because...*) is an NP2-biased verb: participants tend to complete the example fragment with something about Allison. The verb bias works to focus the antecedent in either NP1 or NP2 as a potential antecedent for any anaphor in the second clause, as demonstrated by behavioral experiments using these verbs (Garvey, Caramazza, & Yates 1976; Au 1986; McKoon, Greene, & Ratcliff 1993; Garnham, Traxler, Oakhill, & Gernsbacher 1996; Long & De Ley 2000; Koornneef & Van Berkum 2006).

Van Berkum and colleagues (2007) conducted an ERP experiment to better understand the timing and nature of the processing of coreference in sentences containing verbs of implicit causality. Participants read target sentences within a short discourse context. Pronouns unambiguously coreferred with the antecedent that was consistent (*David apologized to Linda because he was the one to blame*) or inconsistent (*Linda apologized to David because he was not the one to blame*) with the implicit causality bias of the verb. Notice that in the bias-inconsistent case, there is no formal violation at the pronoun; structurally, this pronoun does not constitute a violation. It is only the lexical-semantic feature of implicit causality that may be violated in the bias-inconsistent case. At the critical word, ERPs to bias-inconsistent pronouns were more positive at latencies of 400-700 ms post-stimulus onset, relative to bias-consistent pronouns. This P600 effect for bias inconsistent pronouns indicates that readers were able to make use of a verb's implicit causality information within half a second of reading the pronoun. Additionally, the locus of the effect on the P600 component may inform us about the nature of the processes engaged upon encountering this type of inconsistency. One possibility is that this P600 effect may have arisen because the biasing influence of implicit causality was so strong that readers

tended to take the pronoun as a morphosyntactic error, even when an alternative continuation of the sentence was structurally available. Alternatively, these results might be another example of the P600 effects that have been observed recently in which strong semantic expectations exerted an influence over syntactic processing (Kuperberg 2007).

One way to differentiate between these interpretations is through the use of repeated name coreference. Because repeated names do not carry morphosyntactic information, the observation of a similar P600 effect in response to repeated names that violate the implicit causality bias of a verb would argue against the first possibility presented above. In a recent study (Ledoux, B. Gordon, Roll, & Swaab, 2008), we expanded the design of Van Berkum and colleagues (2007) to include coreferential repeated names in addition to the pronouns. Participants read sentences in which the implicit causality bias of the verb focused the first noun phrase (NP1) or the second noun phrase (NP2); the pronoun or repeated name in the second clause was either consistent with this verb bias (NP1 bias: *Gregory amused Allison because he/Gregory told a very funny joke*; NP2 bias: *Amy thanked Joe because he/Joe had been so helpful*) or inconsistent (NP1 bias: *Gregory amused Allison because she/Allison needed cheering up*; NP2 bias: *Amy thanked Joe because she/Amy was trying to practice good manners*). Additionally, these authors contrasted the focusing effect of implicit causality with that of the discourse prominence of the antecedent: in all of the sentences, the entity in subject position (*Gregory* or *Amy*) was more prominent (according to DPT) than the entity in object position (*Allison* or *Joe*), due to its grammatical role as subject of the sentence. For the pronoun conditions, the results of Van Berkum and colleagues (2007) were replicated: a larger P600 was observed to pronouns that violated the implicit causality bias of the verb, and this was true regardless of the prominence of the antecedent. Interestingly, the locus of the effect of implicit

causality for the repeated name sentences depended on the prominence of the antecedent. When the antecedent of the repeated name was not prominent in the discourse representation (a situation in which repeated name coreference has been shown to be felicitous), the observed effect of implicit causality was similar to that seen for pronouns: the amplitude of the P600 was larger to names that were inconsistent with the bias of the verb. When the antecedent of the repeated name was prominent in the discourse representation, and repeated name coreference was expected to be infelicitous, the effect of verb implicit causality was instead noted on the N400, the amplitude of which was reduced to names that were consistent with the bias of the verb. The interpretation of the P600 effects seen to pronouns and repeated names in these experiments as reflecting the detection of a morphosyntactic error thus seems invalidated, as repeated names do not carry this type of information. An alternative interpretation is that verbs of implicit causality generate such strong semantic expectations that these expectations begin to exert an influence over syntactic processing, as has been suggested for other examples of P600 effects seen to seemingly semantic anomalies (Kuperberg 2007). In the case of repeated names, this was observed only in cases in which semantic integration could proceed without hindrance, that is, in cases in which the repeated name was used felicitously in terms of the establishment of coreference. The infelicitous use of a repeated name to corefer with a prominent antecedent resulted in a disruption of semantic integration processing (similar to that observed previously, Swaab, et al. 2004; Camblin, et al. 2007; Ledoux, et al. 2007), the magnitude of which was influenced by verb consistency.

In summary, we have reviewed in this section studies that did not involve formal violations of the constraints on the establishment of coreference. Instead, these studies involved comparisons of conditions that varied in terms of the relative ease or difficulty of establishing

coreference between a (formally acceptable) anaphor and its antecedent. In some cases, this relative ease or difficulty was a function of the form of the referring expression itself. In others, processing difficulty was determined primarily by manipulations of the context in which the anaphor was found. Still other studies manipulated both of these factors, to determine if the ease or difficulty of processing different referential forms was not absolute, but determined instead by context. In general, the ease or difficulty of establishing coreference with different forms and in different contexts, as reviewed in this section, was reflected in the amplitude of the N400 component of the ERP, with reductions in N400 amplitude seen in cases of relative processing ease. This result confirms and extends the findings of decades of behavioral studies by marking neurophysiologically those conditions in which coreferential processing is easy or difficult. But even more importantly, this result points to a potential mechanism by which “ease” or “difficulty” might be defined. That is, because the effect of manipulations of form of reference or discourse context is reflected in the N400 component, and because this component has been strongly linked to processes of semantic integration, we can (tentatively) conclude that difficulty in semantic integration (and not in syntactic processing) lies at the root of the behavioral effects observed for such accessibility manipulations. The one exception may be in cases of implicit causality, in which the semantic expectation created by verb lexical information may be so strong as to result in a disruption of syntactic processing. The N400 effects observed in these studies have sometimes been accompanied by the differential engagement of working memory processes (as reflected by left anterior negativities) across conditions. Finally, an important new marker of coreferential processing, the Nref, has been found to be useful as an indication of referential ambiguity.

The interaction of lexical and coreferential processes

A handful of ERP studies have compared the processing of words that are used to establish coreference to the expected processing of those same words when coreference is not involved. That is, these studies have examined the modulation of known lexical processing effects by the specific discourse context that is established by coreference. Two factors that have been demonstrated to exert influences on lexical processing (both in word lists and in sentence or discourse contexts) are repetition and semantic or associative relatedness.

Electrophysiologically, the effects of repetition and semantic priming have been linked to a reduction in the amplitude of the N400 to words on their second presentation (Rugg 1985; Smith & Halgren 1987; Besson, Kutas, & Van Petten 1992) or to words that follow semantic associates (Bentin, McCarthy, & Wood 1985; Holcomb & Neville 1990; Holcomb & Anderson 1993; Anderson & Holcomb 1995).

These effects of lexical priming have been contrasted to those exerted by a coreferential discourse context. For example, the inclusion of new names as a lexical match for coreferential repeated names by Ledoux and colleagues (2007) as mentioned above allowed a direct assessment of lexical repetition effects and their modulation by processes of coreference. Similar lexical repetition effects (a reduction of the N400 to repeated relative to new words) were observed in sentences in which the repeated names were used to establish coreference. However, this was only true when the initial name in the sentence (the antecedent for the repeated names) was not prominent in the discourse representation. When the initial name was prominent, no repetition priming effect was observed, suggesting that processes of coreference can sometimes interact with the mechanisms for processing individual words.

Anderson and Holcomb (2005) investigated whether the establishment of coreference between two noun phrases would influence the magnitude of repetition and semantic priming effects during the processing of sentences in a short discourse. An introductory sentence used one of two categorically-related noun phrases to describe an entity in the world (*Kathy sat nervously in the cab/taxi on her way to the airport.*). The next sentence encouraged the establishment of coreference by the use of a definite article, or discouraged it by the use of an indefinite article (*The/A cab came very close to hitting the car*). Lexical repetition effects were assessed by the repetition of the noun phrase across the sentences; lexical semantic priming effects were assessed by the use of semantically related noun phrases (*taxi* and *cab*). Anderson and Holcomb found N400 priming effects at the critical word for both repetitions and synonyms; consistent with previous studies, repetition priming effects were greater than semantic priming effects. The establishment of coreference did not influence the magnitude of these priming effects. Here, then, the establishment of coreference did not interact with more basic mechanisms of word processing. Finally, a larger LAN was observed in the coreferential (definite article) condition relative to the non-coreferential (indefinite article) condition, suggesting the involvement of working memory resources in coreferential processing.

Finally, Ditman, Holcomb, and Kuperberg (2007) examined the interplay of lexico-semantic (categorical) information and discourse context in the process of anaphor resolution. Participants read a short discourse comprising five sentences. The first three sentences introduced different discourse entities. Two of these entities were exemplars of the same category (ex. *A stool is found in a bar* and *A couch is found in a living room*; *stool* and *couch* are exemplars of the category “seats”), while the third introduced an unrelated control (*A clock is found on a wall*). The fourth sentence used the category label itself as an anaphor which referred

back to one of the previously introduced entities; contextual cues in the fourth sentence allowed disambiguation between the two category exemplars. For example, for the sentence *At the bar, Henry sat on the seat*, participants should be able to use the contextual information *at the bar* to comprehend that the seat being referred to here is the stool, and not the couch. The final sentence (*The stool/couch/clock was recently purchased*) reinstated one of the three potential antecedents. At this point in processing, one category exemplar (*stool*) is both congruous with the expected resolution of the anaphor and receives the benefits of lexico-semantic priming from the category label; the other exemplar is incongruous with the expected resolution of the anaphor, but receives lexico-semantic priming benefits; and the control benefits from neither discourse context nor lexico-semantic priming. An examination of the ERP responses to this word allowed an assessment of the relative influence of the coreferential discourse context and lexico-semantic information in the processing of these words. The ERP results suggested an interplay of the two types of information during processing: the amplitude of the N400 component was smallest to the exemplar that benefited from both discourse context and lexico-semantic priming, largest to the control entity (which benefited from neither), and intermediate to the exemplar that benefited from lexico-semantic priming but was incongruous with the discourse context.

In summary, it seems that factors that are seen to affect the processing of words in sentences (such as repetition or semantic relatedness) may sometimes interact with factors that are seen to affect the processing of coreferential relations in determining the amplitude of the N400 component, but this is not always the case. It is difficult at this point to draw strong conclusions about the influence of coreferential discourse context on the processing of individual words that may be subject to other, lexico-semantic influences due to the small number of studies

that have examined this issue. Perhaps future research in this area will allow a determination of the conditions under which the two types of processes are seen to interact.

General conclusions and future directions

Cognitive neuroscientific studies of coreference have provided a number of insights into the neural mechanisms by which anaphors and antecedents are processed and represented. One of the advantages of the event-related potential methodology is the association of different ERP components with different aspects of cognitive processing. This association can then be used to draw tentative inferences about the contribution of different types of information to language comprehension. For example, morphosyntactic information carried by the anaphor is critical to the establishment of coreference. The P600 effect observed to morphosyntactic violations suggests that the processor must engage in additional processes of syntactic structure building or syntactic (re)analysis when the morphosyntactic properties of an anaphor do not match the antecedents available in the local discourse context. But it is not the case that syntactic processing is the only type of processing that is required during the successful establishment of coreference. Instead, the N400 effects observed to manipulations of the availability of the antecedent suggest that processes of semantic integration are also essential during the comprehension of coreferential relations. In fact, the ERP methodology provides us some clues as to the interplay of semantic and syntactic information in coreference. When there is a formal violation of the constraints on coreference, additional syntactic processing must be undertaken, and potential effects of syntactic integration difficulties on the establishment of message-level meaning may be observed downstream from the anaphor (as in the N400 effects seen to the final words of violation sentences). When there is no such formal violation, the syntactic processing

of the anaphor can proceed, and generally a P600 effect is not observed. But message-level meaning in this case can still be differentially affected, even if syntactic processing occurred unheeded, as evidenced by N400 effects at the anaphor in sentences without a formal violation. Thus, the ERP technique offers a unique opportunity to gain greater understanding of the influence of structure versus meaning, and of the interplay of these two types of information, during language processing.

The event-related potential methodology also confers the advantage of high temporal resolution, allowing insight into the temporal dynamics of cognitive processing. This is especially important in the study of language, as most language behavior happens very rapidly, on a time-scale that is difficult to assess with more traditional behavioral techniques. A great contribution of the event-related potential studies discussed above is the demonstration of the immediacy of the engagement of attempts to establish coreference upon encountering a referring expression. The ERP effects observed in studies of coreference typically emerge within half a second of the presentation of the critical stimulus; few other cognitive-behavioral measurement techniques allow the temporal precision that is needed to make such a determination.

In addition to the syntactic and semantic effects observed during coreferential processing, indications of the engagement of working memory resources were observed in several of the ERP studies reviewed above. That working memory would be necessary to the processing of coreference is not surprising; the retrieval of antecedent representations upon encountering the anaphor, the search through various discourse entities for one that matches the anaphor on essential characteristics, and the final selection of one entity as the antecedent of the anaphor are all processes that seem to be defined in terms of working memory. What ERPs are positioned to uniquely contribute in terms of working memory and its role in coreference is an understanding

of the *differential* recruitment of these resources when processing different referring expressions or different types of morphosyntactic violation. Future research might address the interplay of language and other cognitive processes during the establishment of coreference; mechanisms of attention, for example, may prove to be important to our understanding of antecedent accessibility.

Although ERP studies of coreferential processing have begun to draw motivation for their predictions from psycholinguistic processing models, it would seem that there is much room for future studies in this vein. One relatively unexplored avenue is the use of event-related potentials to attempt to differentiate between the predictions offered by the various processing models in ways that behavioral methods have not allowed. Especially useful perhaps would be studies that examine differences in the predictions offered by different models of the *timing* of processing events, as the ERP technique is especially well-suited to the resolution of temporal questions.

To date, few studies have used functional neuroimaging techniques to study the processing of coreference. One reason for the relative paucity of such studies might be the difficulty of the design of discourse materials that are appropriate for presentation in the scanner. Clever designs have overcome this limitation, and it is expected that in the near future, many more will follow suit. Those studies of coreference that have incorporated functional neuroimaging (for example, Hammer et al 2007; Nieuwland, Petersson, & Van Berkum 2007) have tended to support the conclusions drawn from ERP studies, but have offered additional information about the localization of such processing in the brain. The examination of similar questions in conjoined ERP and fMRI studies can offer excellent insight into the temporal mechanisms and spatial representation of language processes.

Short Biography

Kerry Ledoux's research centers around the exploration of the cognitive and neural architectures of language processing. She has focused on the study of psycholinguistics, specifically questions in sentence and discourse processing, using several different research methods, including traditional behavioral measures (reading time, comprehension accuracy, and memory probes during reading), eye movement monitoring, and event-related brain potentials. She has recently published her work in *Memory and Cognition* and in *Psychological Science*. She holds a BA in Psychology from Clark University, Worcester, MA, and an MA and PhD from The University of North Carolina at Chapel Hill. She is currently a post-doctoral researcher at Johns Hopkins University, Baltimore, MD, where she is a member of the Cognitive Neurology/Neuropsychology group.

Chrissy Camblin focuses her research on the effects of context in language processing. Her work has used converging methods of eye tracking and event-related brain potentials (ERPs) to explore the timing and neural bases of language comprehension, as well as functional magnetic brain imaging (fMRI) to examine the brain areas underlying cognitive phenomena. She has recently published her work in the *Journal of Memory and Language* and in *Brain Research*. Chrissy received a BS in Cognitive Science at the University of California at San Diego, and a PhD in Psychological and Brain Science at Duke University. After gaining additional training in cognitive neuroimaging as a post-doctorate researcher at the University of North Carolina at Chapel Hill, she accepted a position at California State University at Monterey Bay where she is currently teaching and developing the college's new psychology major.

References

- Almor, A. 1999. Noun-phrase anaphora and focus: the informational load hypothesis. *Psychological Review* 106. 748-765.
- Almor, A., & V.A. Nair. 2007. The form of referential expressions in discourse. *Language & Linguistics Compass* 1. 84-99.
- Anderson, J.E., & P.J. Holcomb. 1995. Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology* 32. 177-190.
- . 2005. An electrophysiological investigation of the effects of coreference on word repetition and synonymy. *Brain & Language* 94. 200-216.
- Ariel, M. 1988. Referring and accessibility. *Journal of Linguistics* 24. 65-87.
- . 1990. *Accessing noun-phrase antecedents*. London: Routledge.
- Au, T.K.-F. 1986. A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory & Language* 25. 104-122.
- Badecker, W., & K. Straub. 2002. The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 28. 748-769.
- Bentin, S., G. McCarthy, & C.C. Wood. 1985. Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology* 60. 343-355.
- Besson, M., M. Kutas, & C. Van Petten. 1992. An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience* 4. 132-149.
- Brown, C.M., & P. Hagoort. 1993. The processing nature of the N400: evidence from masked priming. *Journal of Cognitive Neuroscience* 5. 34-44.
- Camblin, C.C., K. Ledoux, M. Boudewyn, P.C. Gordon, & T.Y. Swaab. 2007. Processing new and repeated names: effects of coreference on repetition priming with speech and fast RSVP. *Brain Research* 1146. 172-184.
- Chambers, C.G., & R. Smyth. 1998. Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory & Language* 39. 593-608.
- Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht: Foris.

- Clark, H.H., & C.J. Sengul. 1979. In search of referents for nouns and pronouns. *Memory & Cognition* 7. 35-41.
- Coulson, S., J.W. King, & M. Kutas. 1998. Expect the unexpected: event-related brain response to morphosyntactic violations. *Language & Cognitive Processes* 13. 21-58.
- Daneman, M., & P.A. Carpenter. 1980. Individual difference in working memory and reading. *Journal of Verbal Learning & Verbal Behavior* 19. 450-466.
- Ditman, T., P.J. Holcomb, & G.R. Kuperberg. 2007. The contributions of lexico-semantic and discourse information to the resolution of ambiguous categorical anaphors. *Language & Cognitive Processes* 22. 793-827.
- Ehrlich, K. 1980. Comprehension of pronouns. *Journal of Experimental Psychology* 32. 247-255.
- Ehrlich, K., & K. Rayner. 1983. Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning & Verbal Behavior* 22. 75-87.
- Filik, R., A.J. Sanford, & H. Leuthold. 2007. Processing pronouns without antecedents: evidence from event-related brain potentials. Poster presented at the 14th Annual Cognitive Neuroscience Society Meeting, New York, NY.
- Friederici, A.D. 1995. The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain & Language* 50. 259-281.
- . 1998. Diagnosis and reanalysis: two processing steps the brain may differentiate. *Reanalysis in sentence processing*, ed. by J.D. Fodor & F. Ferreira, 177-200. Dordrecht: Kluwer Academic Publishers.
- . 2001. Syntactic, prosodic, and semantic processes in the brain: evidence from event-related neuroimaging. *Journal of Psycholinguistic Research* 30. 237-250.
- . 2002. Towards a neural basis of auditory sentence processing. *TRENDS in Cognitive Sciences* 6. 78-84.
- Garnham, A. 2001. *Mental models and the interpretation of anaphora*. Philadelphia PA: Taylor and Francis.
- Garnham, A. & J. Oakhill. 1985. On-line resolution of anaphoric pronouns: Effects of inference making and verb semantics. *British Journal of Psychology* 76. 385-393.
- Garnham, A., M. Traxler, J. Oakhill, & M.A. Gernsbacher. 1996. The locus of implicit causality effects in comprehension. *Journal of Memory & Language* 35. 517-543.

- Garrod, S., & A. Sanford. 1977. Interpreting anaphoric relations: the integration of semantic information while reading. *Journal of Verbal Learning & Verbal Behavior* 16. 77-90.
- Garvey, C., A. Caramazza, & J. Yates. 1976. Factors influencing assignment of pronoun antecedents. *Cognition* 3. 227-243.
- Gernsbacher, M.A. 1990. *Language comprehension as structure building*. Hillsdale, NJ: Erlbaum.
- Gordon, P.C., B.J. Grosz, & L.A. Gilliom. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17. 311-347.
- Gordon, P.C., & R. Hendrick. 1998. The representation and processing of coreference in discourse. *Cognitive Science* 22. 389-424.
- Gordon, P.C., R. Hendrick, K. Ledoux, & C.-L. Yang. 1999. Processing of reference and the structure of language: An analysis of complex noun phrases. *Language & Cognitive Processes* 14. 353-379.
- Gordon, P.C., & K.A. Searce. 1995. Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory & Cognition* 23. 313-323.
- Greene, S.B., G. McKoon, & R. Ratcliff. 1992. Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 18. 266-283.
- Grober, E.H., W. Beardsley, & A. Caramazza. 1978. Parallel function strategy in pronoun assignment. *Cognition* 6. 117-133.
- Grosz, B.J., A.K. Joshi, & S. Weinstein. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21. 203-226.
- Hagoort, P. 2003. How the brain solves the binding problem for language: A neurocomputational model of syntactic processing. *NeuroImage* 20. S18-S29.
- Hagoort, P., C. Brown, & J. Groothusen. 1993. The syntactic positive shift as an ERP-measure of syntactic processing. *Language & Cognitive Processes* 8. 439-483.
- Hammer, A., R. Goebel, J. Schwarzbach, T.F. Münte, & B.M. Jansma. 2007. When sex meets syntactic gender on a neural basis during pronoun processing. *Brain Research* 1146. 185-198.
- Hammer, A., B.M. Jansma, M. Lamers, & T.F. Münte. 2005. Pronominal reference in sentences about persons or things: An electrophysiological approach. *Journal of Cognitive Neuroscience* 17. 227-239.

- Harris, T., K. Wexler, & P. Holcomb. 2000. An ERP investigation of binding and coreference. *Brain & Language* 75. 313-346.
- Holcomb, P.J. 1993. Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. *Psychophysiology* 30. 47-61.
- Holcomb, P.J., & J.E. Anderson. 1993. Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language & Cognitive Processes* 8. 379-411.
- Holcomb, P.J., & H.J. Neville. 1990. Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language & Cognitive Processes* 5. 281-312.
- Kemmerer, D., C. Weber-Fox, K. Price, C. Zdanczyk, & H. Way. 2006. *Big brown dog or brown big dog?* An electrophysiological study of semantic constraints on prenominal adjective order. *Brain & Language* 100. 238-256.
- Kim, A., & L. Osterhout. 2005. The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory & Language* 52. 205-225.
- King, J.W., & 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. 376-395.
- Kluender, R., & M. Kutas. 1993. Bridging the gap: evidence from ERPs on the processing of unbound dependencies. *Journal of Cognitive Neuroscience* 5. 196-214.
- Kolk, H. & D. Chwilla. 2007. Late positivities in unusual situations. *Brain & Language* 100. 257-261.
- Koorneef, A.W., & J.J.A. Van Berkum. 2006. On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory & Language* 54. 445-465.
- Kuperberg, G.R. 2007. Neural mechanisms of language comprehension: challenges to syntax. *Brain Research* 1146. 23-49.
- Kuperberg, G.R., D.A. Kreher, T. Sitnikova, D.N. Caplan, & P.J. Holcomb. 2006. The role of animacy and thematic relationships in processing active English sentences: evidence from event-related potentials. *Brain & Language* 100. 223-237.
- Kutas, M., & S.A. Hillyard. 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207. 203-205.
- Ledoux, K., B. Gordon, M. Roll, & T.Y. Swaab. 2008. The relative influence of verb implicit causality and discourse prominence on the establishment of coreference with repeated

- names and pronouns: An ERP study. Poster presented at the 21st Annual CUNY Conference on Human Sentence Processing, Chapel Hill, NC.
- Ledoux, K., P.C. Gordon, C.C. Camblin, & T.Y. Swaab. 2007. Coreference and lexical repetition: mechanisms of discourse integration. *Memory & Cognition* 35. 801-815.
- Long, D.L., & L. De Ley. 2000. Implicit causality and discourse focus: The interaction of text and reader characteristics in pronoun resolution. *Journal of Memory & Language* 42. 545-570.
- McDonald, J., & B. MacWhinney. 1995. The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory & Language* 34. 543-566.
- McKoon, G., S.B. Greene, & R. Ratcliff. 1993. Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology: Learning, Memory & Cognition* 19. 1040-1052.
- Nair, V.A., A. Almor, & J.M.C. Vendemia. 2007. ERP study of anaphor resolution in sentences and discourse. Poster presented at the 14th Annual Cognitive Neuroscience Society Meeting, New York, NY.
- Nieuwland, M.S., M. Otten, & J.J.A. Van Berkum. 2007. Who are you talking about? Tracking discourse-level referential processing with event-related brain potentials. *Journal of Cognitive Neuroscience* 19. 228-236.
- Nieuwland, M.S., K.M. Petersson, & J.J.A. Van Berkum. 2007. On sense and reference: Examining the functional neuroanatomy of referential processing. *Neuroimage* 37. 993-1004.
- Nieuwland, M.S., & J.J.A. Van Berkum. 2006. Individual differences and contextual bias in pronoun resolution: evidence from ERPs. *Brain Research* 1118. 155-167.
- Osterhout, L. 1997. On the brain response to syntactic anomalies: manipulations of word position and word class reveal individual differences. *Brain & Language* 59. 494-522.
- Osterhout, L., M. Bersick, & J. McLaughlin. 1997. Brain potentials reflect violations of gender stereotypes. *Memory & Cognition* 25. 273-285.
- Osterhout, L., & P.J. Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31. 785-806.
- Osterhout, L., & L.A. Mobley. 1995. Event-related brain potentials elicited by failure to agree. *Journal of Memory & Language* 34. 739-773.
- Rugg, M.D. 1985. The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology* 22. 642-647.

- Sanford, A.J., & S.C. Garrod. 1998. The role of scenario mapping in text comprehension. *Discourse Processes* 26. 159-190.
- Schmitt, B.M., M. Lamers, & T.F. Münte. 2002. Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. *Cognitive Brain Research* 14. 333-346.
- Sheldon, A. 1974. The role of parallel function in the acquisition of relative clauses in English. *Journal of Verbal Learning & Verbal Behavior* 13. 272-281.
- Smith, M.E., & E. Halgren. 1987. Event-related potentials during lexical decision: effects of repetition, word frequency, pronounceability, and concreteness. *Electroencephalography and Clinical Neurophysiology Suppl.* 40. 417-421.
- Smyth, R.H. 1994. Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research* 23. 197-229.
- Streb, J., E. Hennighausen, & F. Rösler. 2004. Different anaphoric expressions are investigated by event-related brain potentials. *Journal of Psycholinguistic Research* 33. 175-201.
- Streb, J., F. Rösler, & E. Hennighausen. 1999. Event-related responses to pronoun and proper name anaphors in parallel and non-parallel discourse structures. *Brain & Language* 70. 273-286.
- Swaab, T.Y., C.C. Camblin, & P.C. Gordon. 2004. Reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience* 16. 715-726.
- Van Berkum, J.J.A., C.M. Brown, & P. Hagoort. 1999. Early referential context effects in sentence processing: evidence from event-related brain potentials. *Journal of Memory & Language* 41. 147-182.
- Van Berkum, J.J.A., C.M. Brown, P. Hagoort, & P. Zwitserlood. 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology* 40. 235-248.
- Van Berkum, J.J.A., A.W. Koornneef, M. Otten, & M.S. Nieuwland. 2007. Establishing reference in language comprehension: an electrophysiological perspective. *Brain Research* 1146. 158-171.
- Van Berkum, J.J.A., P. Zwitserlood, M. Bastiaansen, C. Brown, & P. Hagoort. 2004. So who's "he" anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension. Poster presented at the 11th Annual Cognitive Neuroscience Society Meeting, San Francisco, CA.

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

van Herten, M., H.H. Kolk, & D.J. Chwilla. 2005. An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research* 22. 241-255.

Vonk, W. 1984. Eye movements during the comprehension of pronouns. *Theoretical and applied aspects of eye movement research*, ed. by A.G. Gale & F. Johnson, 203-212. Amsterdam: North-Holland.

---. 1985. On the purpose of reading and the immediacy of processing pronouns. *Eye movements and human information processing*, ed. by R. Groner, G.W. McConkie, & C. Menz, 207-215. Amsterdam: Elsevier.