

# Sensitivity to Referential Ambiguity in Discourse: The Role of Attention, Working Memory, and Verbal Ability

Megan A. Boudewyn, Debra L. Long, Matthew J. Traxler, Tyler A. Lesh, Shruti Dave, George R. Mangun, Cameron S. Carter, and Tamara Y. Swaab

## Abstract

■ The establishment of reference is essential to language comprehension. The goal of this study was to examine listeners' sensitivity to referential ambiguity as a function of individual variation in attention, working memory capacity, and verbal ability. Participants listened to 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). The manipulation rendered an anaphor in a subsequent sentence (e.g., oak) ambiguous or unambiguous. EEG was recorded as listeners comprehended the story, after which participants completed tasks to assess working memory, verbal ability, and the ability to use context in task performance. Power in the alpha and theta frequency bands when listeners received critical information about the discourse entities (e.g., oaks) was used to index attention and the

involvement of the working memory system in processing the entities. These measures were then used to predict an ERP component that is sensitive to referential ambiguity, the Nref, which was recorded when listeners received the anaphor. Nref amplitude at the anaphor was predicted by alpha power during the earlier critical sentence: Individuals with increased alpha power in ambiguous compared with unambiguous stories were less sensitive to the anaphor's ambiguity. Verbal ability was also predictive of greater sensitivity to referential ambiguity. Finally, increased theta power in the ambiguous compared with unambiguous condition was associated with higher working-memory span. These results highlight the role of attention and working memory in referential processing during listening comprehension. ■

## INTRODUCTION

The goal of language comprehension is to construct an accurate and coherent representation of a speaker's or author's intended meaning; this mental representation is often called a discourse model. Comprehension involves language-specific processes, such as phonological/orthographic analysis and syntactic parsing, and domain-general ones, such as focusing attention, retrieving, and maintaining discourse-relevant information in memory. Of particular interest in this study are processes that are related to sustaining attention to task-relevant input and those that are related to the activation and maintenance of information in working memory (WM). Sustained attention is critical in language comprehension because linguistic input is sequential and often occurs over long periods. Similarly, the WM system is essential to comprehension; readers and listeners must maintain recently processed information and information that is activated from long-term memory to understand incoming words and sentences.

In this study, we examined the roles of attention and WM processes in the establishment of discourse reference. Reference is established when readers or listeners recognize that an incoming word, an anaphor, refers to

the same entity as a word that was introduced previously, the antecedent. Failure to establish reference has significant negative consequences on comprehension. An anaphor like "she" has limited meaning, unless it is linked to an antecedent (i.e., who is she?). Likewise, an anaphor like "oak" can be ambiguous if two "oaks" have been introduced in the preceding context. The goal of this study was to examine how referential processing is affected by fluctuations in attention to linguistic input and by variation in demands on WM.

We hypothesized that both attention and WM processes are important in comprehending referential expressions. With respect to attention, we hypothesize that lapses of attention during key points in the construction of a discourse model should result in subsequent failure to make use of the critical information. For example, inattention to the input when an entity is introduced should result in later comprehension failure if that entity becomes the antecedent of an anaphor. With respect to WM, we hypothesize that encoding and maintenance processes in WM should be involved in incorporating an antecedent into the discourse model and in later accessing that information when an anaphor is encountered.

Here, we examined oscillatory activity in the alpha and theta frequency bands as neural correlates of variation in attention and WM processes, respectively. Although

recent electrophysiology (EEG) studies have examined how alpha and theta oscillations correlate with attention and WM processes, few of these studies have investigated oscillations in these frequency bands during language processing. Language comprehension is an interesting context in which to investigate attention and WM processes for two reasons. First, linguistic input is received sequentially and over considerable lengths of time. Numerous studies have shown that attention waxes and wanes during language processing and that lapses of attention are associated with impaired comprehension (e.g., Franklin, Broadway, Mrazek, Smallwood, & Schooler, 2013; Franklin, Smallwood, & Schooler, 2011; Reichle, Reineberg, & Schooler, 2010; Smallwood, McSpadden, & Schooler, 2008; Smallwood, Fishman, & Schooler, 2007). Second, language comprehension places considerable storage and processing demands on a limited-capacity WM system. The relation between WM capacity and comprehension skill is well documented (see Long, Johns, & Morris, 2006, for a review).

To investigate the role of attention and WM processes during discourse comprehension, we measured ERPs to anaphors in the final sentence of short spoken stories. We examined whether the ERPs to these expressions were correlated with EEG activity that occurred earlier in the stories, namely, when their antecedents were first introduced. We also investigated whether these EEG measures were related in a meaningful way to individual differences in behavioral measures of WM capacity. In the following sections, we describe the Nref, an ERP effect that is sensitive to the processing of referential expressions during comprehension. We then describe previous research that has connected oscillations in the alpha and theta bands to attention, WM, and language processing. Finally, we describe our experimental manipulation and hypothesized results.

### **The Nref Effect**

We examined referential processing using an ERP measure of referential ambiguity called the Nref effect. The Nref is a sustained frontal negative deflection elicited when incoming written or spoken words are referentially ambiguous, that is, when more than one antecedent is possible in the preceding context (Nieuwland & Van Berkum, 2006, 2008; Van Berkum, 2008; Nieuwland, Otten, & Van Berkum, 2007; Van Berkum, Brown, Hagoort, & Zwitserlood, 2003; Van Berkum, Brown, & Hagoort, 1999; see Van Berkum, Koornneef, Otten, & Nieuwland, 2007, for a review). The Nref has been observed when readers/listeners encounter an anaphor, such as “the lecturer,” in a story in which two antecedents have been introduced in the previous context, as compared with when only one has been mentioned (e.g., two lecturers vs. one lecturer; e.g., Van Berkum et al., 1999, 2003). Importantly, the Nref tracks referential ambiguity at the message level: It is not sufficient that two possible antecedents have been mentioned; both must be available in the developing represen-

tation of the context (Nieuwland et al., 2007). For example, if the context introduces two lecturers but then states that one of them left the room, a subsequent reference to the lecturer is not referentially ambiguous and does not elicit an Nref (Nieuwland et al., 2007).

These properties make the Nref useful for examining a listener’s ability to accurately encode information about entities in the discourse context and to retrieve this information to resolve referential ambiguities. As noted above, establishing reference is essential in discourse comprehension because it links current input to preceding discourse (e.g., which of the two possible lecturers is a speaker talking about?). Importantly, the establishment of reference depends on having accurate information about potential antecedents in memory. If listeners were not attending to the input when two lecturers were introduced, they will be insensitive to the ambiguity of the anaphor lecturer at a later point in the story. In addition, WM processes should be involved in tracking available antecedents over the course of a story.

### **Oscillatory Activity and Its Relation to Attention, WM, and Language Processing**

Changes in oscillatory activity in the alpha band (~8–12 Hz) have long been associated with changes in attentional focus (Klimesch, 1999, 2012; Mazaheri et al., 2010; Sauseng, Klimesch, Doppelmayr, et al., 2005; Sauseng, Klimesch, Schabus, & Doppelmayr, 2005; Gevins, Smith, McEvoy, & Yu, 1997; Pfurtscheller, Stancak, & Neuper, 1996; Van Winsum, Sergeant, & Geuze, 1984; Adrian & Matthews, 1934). Two lines of research suggest that alpha power is sensitive to attention during the performance of complex cognitive tasks. One line of research has examined changes in alpha during visual attention tasks (Bengson, Mangun, & Mazaheri, 2012; Thut, Nietzel, Brandt, & Pascual-Leone, 2006). In these tasks, attention to task-relevant information requires directing attention to external stimuli. For example, several studies have shown that increases in alpha power precede errors on tasks such as the sustained attention to response task, in which participants must withhold a prepotent response on no-go trials; errors on such trials are often associated with lapses in attention (O’Connell, Dockree, Bellgrove, et al., 2009; O’Connell, Dockree, Robertson, et al., 2009; Dockree, Kelly, Foxe, Reilly, & Robertson, 2007; Dockree et al., 2004; Roche et al., 2004). Other studies have shown that changes in alpha correlate with attention to specific locations in visual space. Specifically, decreases in alpha power occur contralateral to a to-be-attended location (i.e., the location at which the target will appear), and increases in alpha power occur contralateral to the to-be-ignored location (Bengson et al., 2012; Thut et al., 2006). Studies such as these suggest that reductions in alpha power correlate with attention that is directed toward processing stimuli from the external environment.

Another line of research has examined changes in alpha in WM tasks. In a WM task, attention must be directed to

an external stimulus during encoding and then shift to an internal representation of the stimulus during the maintenance period. In these studies, alpha power increases as WM load increases (Meyer, Obleser, & Friederici, 2013; Jensen, Gelfand, Kounios, & Lisman, 2002; Röhms, Klimesch, Haider, & Doppelmayr, 2001). For example, Jensen et al. (2002) found that alpha power increased as a function of the number of items held in WM during the delay between presentation of a visual array and a recognition test. The results from studies of alpha activity in WM tasks may appear to be in conflict with those from studies of alpha activity in visual attention tasks, if changes in alpha are taken to reflect changes in attention to task-relevant information. However, the two lines of research are complementary if changes in alpha reflect changes in attention to external stimuli; whether external information is task relevant depends on the task (as also suggested by Roux & Uhlhaas, 2014; Mazaheri et al., 2010). Increases in alpha are observed when input from the environment is ignored, suppressed, or unattended, which is consistent with accounts that ascribe inhibition-related functions to oscillations in the alpha frequency band (Klimesch, 2012; Jensen & Mazaheri, 2010). Inhibition of external input is advantageous in WM tasks in which task-relevant processing requires attention to internal representations, whereas inhibition of external input is disadvantageous in tasks in which attention is directed to external stimuli, as in visual attention tasks. In the current study, we use alpha power as an index of attention to linguistic input in a story listening task.

Oscillations in the theta frequency range (~4–7 Hz) have been linked to memory processes in a number of studies that have investigated scalp-recorded EEG/MEG in humans (Hsieh, Ekstrom, & Ranganath, 2011; Sauseng, Griesmayr, Freunberger, & Klimesch, 2010; Meltzer et al., 2008; Jensen & Tesche, 2002; Gevins & Smith, 2000; Gevins et al., 1997). In particular, increases in theta oscillations have been associated with increased WM demands. In one study, participants were asked to remember lists of digits over a 3-sec delay; the delay period elicited activity in the theta band that increased along with the number of digits to be maintained (Jensen & Tesche, 2002). Other studies have also found that theta power increases as a function of increases in WM load and task difficulty (Meltzer et al., 2008; Gevins et al., 1997). Furthermore, Hsieh and colleagues (2011) reported results suggesting that theta oscillations may be specifically related to WM tasks that require the maintenance of temporal order information. This latter finding has important implications for the role of theta in language processing, as language comprehension requires that listeners/readers track sequences of events to construct a coherent discourse model. Indeed, studies of sentence processing have shown that theta power increases as a sentence unfolds and in response to sentences that place demands on semantic-memory retrieval processes (Hald, Bastiaansen, & Hagoort, 2006; Bastiaansen, Van Berkum, & Hagoort, 2002a, 2002b). On

the basis of these findings, theta power was examined in this study as a correlate of WM processing to investigate how demands on WM affect referential processing during comprehension.

## The Current Study

Our goal in this study was to examine the roles of attention and WM processes in the establishment of discourse reference. Participants listened to short stories and answered true/false comprehension questions after each one. EEG activity was recorded as the stories were presented. Each story was four sentences long, and the fourth sentence (S-4) always contained an anaphor. We manipulated the referential connection between the entities introduced in S-3 and the anaphor in S-4 as can be seen in the following example:

- S-1: A lumberjack hiked into a forest carrying a chainsaw.  
S-2: He was going to cut down a tree.  
S-3: Unambiguous/ambiguous: In a clearing he found an **oak** that had a mushroom on it, and an **elm/oak** that had birds in its branches.  
S-4: The lumberjack cut down the **OAK**...  
True or False: The lumberjack was unable to find a clearing.

In the unambiguous condition, two entities were introduced in S-3 (an oak with a mushroom and an elm with birds), but only one entity was a possible antecedent of the anaphor in S-4 (the oak). In the ambiguous condition, two entities were introduced (e.g., an oak with a mushroom and an oak with birds), and both were possible antecedents of the subsequent anaphor. We predicted an Nref effect in the group as a whole for ERPs that were time locked to the anaphor in S-4, such that a larger frontal negativity would be found in the ambiguous than unambiguous condition, reflecting listeners' understanding that the anaphor was ambiguous when two potential antecedents were available.

In addition to the EEG data, we collected behavioral measures of individual variation in WM capacity: performance on the listening span and operation span tasks. Nieuwland and Van Berkum (2006) have shown that the Nref effect is correlated with individual differences in WM span. Thus, we expected a positive correlation between WM span and the Nref: Listeners who are high in WM span should be more sensitive to the referential ambiguity than listeners who are low in span. We also included two other behavioral measures: performance on the AX version of the continuous performance task (AX-CPT) and performance on a test of vocabulary size (Nelson–Denny Vocabulary). In the AX-CPT, participants respond to a target based on the identity of a previous stimulus. It has been used in previous studies to examine individual variation in the ability to maintain and use context in the performance of a task (Cohen, Barch, Carter,

& Servan-Schreiber, 1999). The ability to use context is essential in the processing of anaphors; thus, performance on the AX-CPT may be predictive of the Nref effect. We included a measure of vocabulary size because it has been shown that it accounts for more variation in language comprehension ability than do individual differences in other linguistic and cognitive abilities (Hamilton, Freed, & Long, 2013; Blozis & Traxler, 2007; Long et al., 2006). We used a multiple regression approach to determine whether performance on the span tasks, the AX-CPT, and the vocabulary test was unique in predicting variation in the Nref effect.

With respect to the relation between oscillatory activity and referential processing, we hypothesized that power in the alpha frequency band, measured when entities were introduced in S-3, would predict the size of the Nref when listeners processed the anaphor in S-4. We calculated the difference between alpha power when listeners processed S-3 in the ambiguous and unambiguous conditions. We then correlated this difference with the size of the Nref effect. Listeners who were attending to the critical information in S-3 should have decreased alpha power relative to listeners who were not attending. Furthermore, these same listeners should show a larger Nref effect. To ensure that the correlation between alpha power at S-3 and the Nref effect at S-4 was specific to attention during processing of the critical input in S-3, we also calculated the difference between alpha power across conditions when listeners processed S-2. We predicted that the size of the Nref effect should be related to differences in alpha power at S-3, when critical information was presented, but not to differences in power at S-2.

We also hypothesized that theta power during S-3, when the antecedents were introduced, would predict the size of the Nref effect at the anaphor in S-4. When entities are introduced in discourse, they must be encoded and integrated into the discourse model. The conceptual similarity of the entities may affect how easy or difficult this is to do. The greater the similarity between two entities, the more difficult it may be to create separate representations of them. If theta power is sensitive to sentence processing demands, then power in this frequency band should be greater in the ambiguous condition than in the unambiguous condition. Moreover, these power differences may be related to later processing of the anaphor. In particular, listeners who show a greater power difference at S-3, when the entities are introduced, may show a larger Nref effect.

## METHODS

### Participants

Thirty-five participants (nine men) gave informed consent and participated in this study, which was approved by the University of California Davis' institutional review

board. All were right-handed, native speakers of English, with no reported problems with hearing or reading nor any neurological/psychological disorders. All were compensated with course credit. The average age of participants was 19.9 (range = 18–27) years.

### Materials

One hundred sixty experimental stories were constructed, each with two versions (see above for an example story). Each story consisted of four sentences and was written such that information introduced in S-3 would be critical to understanding S-4. Namely, S-3 introduced two entities, one of which was then referred to by an anaphor in S-4. The entities were either distinct (e.g., *an oak and an elm*) or were very similar (an oak and another oak). The anaphor in S-4 was unambiguous or ambiguous depending on the manipulation in S-3. The anaphor was unambiguous when the entities were distinct (oak/elm) and ambiguous when they were very similar (e.g., oak/oak). ERPs were time locked to the anaphor in the final sentence (S-4). Critical target words (oak) never appeared in the sentence-final position.<sup>1</sup> Stories were divided into lists and counterbalanced such that participants were presented with equal numbers of stories in each condition but only one condition from each story set (80 stories per condition [160 in total], per list). An additional 115 filler stories were included in each list. Filler stories were designed to be very similar in form to the experimental ones, consisting of four-sentence narratives. About half of the filler stories introduced multiple antecedents; these were presented in S-1 or S-2 but not S-3.

All stories were digitally recorded using a Schoeps MK2 microphone and a Sound Devices USBPre A/D (Reedsburg, WI; 44,100 Hz, 16 bit), Reedsburg, WI and were spoken by a woman, with natural inflection and at a natural speaking rate. Acoustic onset and offset of the anaphor were determined by visual inspection of the speech waveform and by listening to the words using speech editing software (Audacity; audacityteam.org). Discourse context (the first three sentences of each story) and the final sentences were recorded separately. The average duration of the anaphor was 462.2 msec (range = 254–834 msec). The duration of the anaphors did not differ across conditions ( $p = .45$ ). Stories were followed by comprehension questions; the correct response to half of the questions was true and to the other half was false (see above for an example question).

### Procedure

Participants were tested in two sessions, no more than a week apart. In Session 1, participants completed the EEG portion of the study. Participants were seated in a comfortable chair in an electronically shielded, sound-attenuating booth. Stimuli were presented through Beyer (Farmingdale, NY) dynamic headphones using Presentation software

(neurobs.com). Trials began with a white fixation cross presented in the center of a black screen, about 100 cm in front of the participants. The fixation cross was present for 1000 msec before the onset of each story, throughout story presentation, and for 1000 msec after story offset. It was then replaced by a comprehension question, which remained on the screen until a response was made. Participants were asked to make a true/false response by pressing corresponding keys on a keyboard. Sessions were divided into 10 blocks for presentation to allow for short breaks. The order of blocks and lists was counter-balanced. Participants were asked to maintain center fixation on the cross whenever it was present to minimize eye movement artifacts in the signal but were not required to refrain from blinking throughout story presentation (given the relatively long trial duration of four spoken sentences per story). Participants were additionally instructed that they were free to make eye movements or to rest their eyes in between stories, during the comprehension question portion of each trial.

During Session 2, participants completed a listening span task, an operation span task, the AX-CPT, and the Nelson–Denny vocabulary and comprehension tests. The listening span task was adapted from Daneman and Carpenter (1980) and consisted of 25 sets of sentences ranging from two sentences per set to six; there were five sets of each set size. Participants were instructed to listen to all sentences within each set for comprehension and to indicate whether each sentence was true or false immediately after hearing it. In addition, participants were instructed to remember the final word of each sentence in the set and were asked to recall them in any order after the whole set was presented. There was a 1500-msec pause between each sentence during which participants made their true/false response. Presentation of sets was random. Each correct response (correct recall of the final word) was scored as one point, for a maximum of 100 points. The task, which was adapted from its visual counterpart (reading span), predicts reading comprehension and syntactic parsing abilities, particularly when used in conjunction with other WM tasks (Waters & Caplan, 2003). Participants also completed a second measure of WM capacity. The operation span task was adapted from Unsworth, Redick, Heitz, Broadway, and Engle (2009) and consisted of 15 sets of equation/letter pairs, ranging from three pairs per set to seven; there were three sets of each set size. Participants were instructed to complete each math equation and to remember the letter that appeared afterward. After each set of equation/letter pairs was presented, participants were asked to recall, in order, the letters that had appeared in the set. To prevent rehearsal of the letters, a limit was imposed on the time allowed to solve the math equation. The time limit consisted of the average time to solve the problem, determined in a set of practice trials, plus 2.5 *SDs*. After this amount of time, the program automatically moved to the next equation. Span was calculated as the number of cor-

rectly recalled words for those equations that were correctly solved, for a maximum of 75 points.

The AX-CPT was designed to assess cognitive control mechanisms, specifically the ability to maintain task goals and task-relevant context (Cohen et al., 1999). The version used here was developed as part of the Cognitive Neuroscience Test Reliability and Clinical Applications for Schizophrenia initiative and is freely available via the Cognitive Neuroscience Test Reliability and Clinical Applications for Schizophrenia Web site (<http://cntracs.ucdavis.edu/task/dpx>). In the AX-CPT, participants respond to a series of letters by pressing a key with their index finger or pressing a different key with their middle finger depending on the trial type. They are told to respond with their index finger only to the letter X (target) and only if the X is preceded by the letter A (cue) and to respond with their middle finger to all other letters. Cue-target sequences are frequent (70%) and set up a tendency to make a target response when the letter X occurs. As a result, participants often make errors when the letter X appears but was not preceded by an A (e.g., trials in which an X is preceded by the letter B). Good performance on the task involves the ability to attend to the context in which the letter X occurs and to maintain the task goal throughout processing.

We administered the vocabulary section of the Nelson–Denny Reading Test, a standardized paper-and-pencil assessment. Participants were allotted 15 min to complete the section. Vocabulary is strongly predictive of language comprehension abilities (Long, Prat, Johns, Morris, & Jonathan, 2008) and was employed in this study to examine individual variation in the ERP data that may be because of variability in verbal ability. Nelson–Denny vocabulary scores were calculated as the total correct for the vocabulary section, Form G (out of 80 possible).

EEG was recorded from 29 tin electrodes, mounted in a custom elastic cap (ElectroCap International, Eaton, OH). The right mastoid electrode served as the recording reference (except for four electrodes used to measure eye movements: One electrode above and one below the left eye were referenced to each other, and two placed on the outer canthi of each eye were referenced to each other). The left mastoid electrode was used for later algebraic rereferencing (to the average of both mastoids). The EEG signal was amplified with bandpass cutoffs at 0.05 and 30 Hz and digitized at a sampling rate of 250 Hz (Neuroscan Synamps II, Charlotte, NC). EEG was digitized continuously with stimulus codes that were used for subsequent time locking and averaging. Impedances were kept below 5 k $\Omega$ . Data processing and analysis was performed using SCAN (Compumedics Neuroscan) and Matlab (The MathWorks, Natick, MA), using the EEGLAB toolbox and ERPLab plugin, with custom Matlab and UNIX routines. Independent components analysis artifact correction was used to correct for eye blinks.

To calculate oscillatory measures at S-3, single-trial waveforms were screened for amplifier blocking, muscle artifacts, and horizontal eye movements over epochs of 4000 msec, starting 200 msec before the onset of S-3.

This epoch covered all or most of the duration of S-3 (average = 5008.5 msec, range = 3022–7505 msec). This duration was chosen so as to include the presentation of both entities (oak/elm or oak/oak) in S-3. Artifact rejection procedures resulted in the rejection of 19.3% of trials, on average. EEG spectral power was calculated using the EEGlab toolbox by convolving single-trial epochs in the unambiguous and ambiguous conditions with seven-cycle complex Morlet wavelets. Power for 52 log-spaced frequencies from 4 to 30 Hz was averaged across trials in each of the two conditions and log transformed with respect to EEG activity in the 200 msec preceding the onset of S-3. Power estimates for each frequency were binned into theta (5–7 Hz) and alpha (9–12 Hz) frequency bands. These bands were selected to maximize our ability to test for differences between bands. Following the approach used by Hsieh et al. (2011), relatively narrow, spaced bands were selected to minimize overlap among bands and the impact of frequency smearing from time–frequency decomposition. In the analyses presented below, activity in each band for the unambiguous condition was subtracted from activity in the ambiguous condition, which yielded an ambiguous > unambiguous difference measure. Electrodes were grouped into seven clusters for analysis: left frontal (F3, F7, and FC5), midfrontal (AFz and Fz), right frontal (F4, F8, and FC6), central (FC1, FC2, Cz, CP1, and CP2), left posterior (CP5, P3, and T5), midposterior (Pz and POz), and right posterior (CP6, P4, and T6).

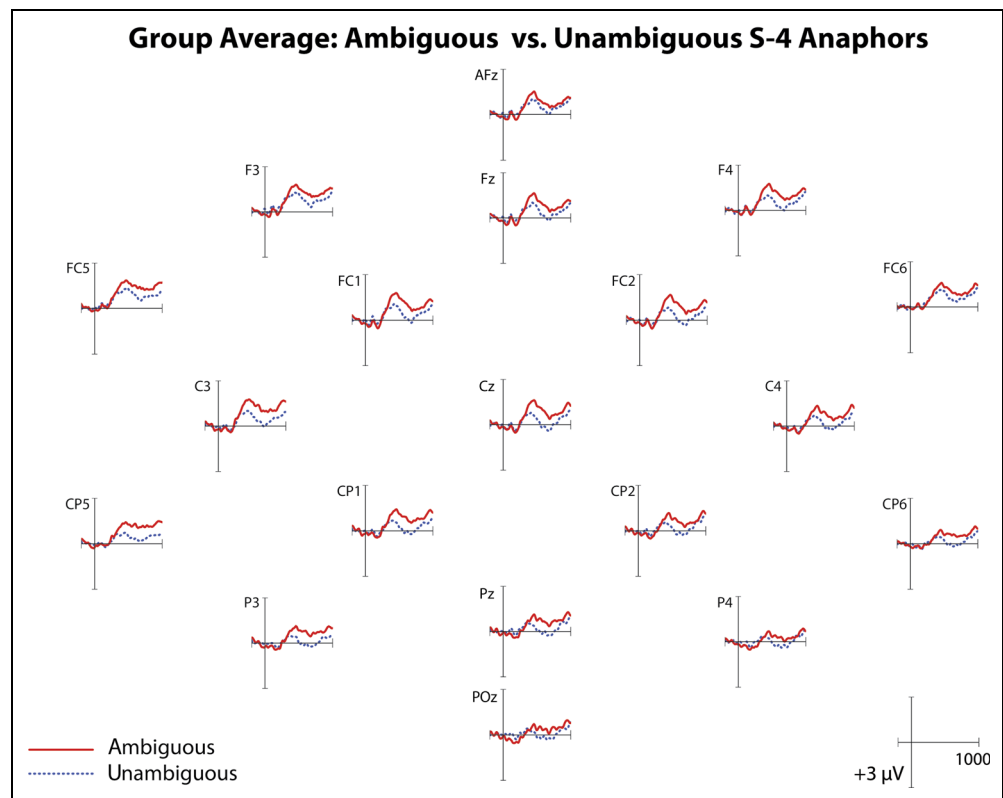
To calculate the Nref effect to the anaphors, single-trial waveforms were screened for amplifier blocking, muscle artifacts, and horizontal eye movements over epochs of 1200 msec, starting 200 msec before the onset of the anaphors in S-4. This resulted in the rejection of 8.9% of trials. Average ERPs that were time locked to the anaphor (ambiguous/unambiguous) were computed over artifact-free trials and were filtered off-line with a Gaussian low-pass filter (25-Hz half-amplitude cutoff).

Three participants were excluded from all analyses because too few uncontaminated trials remained after artifact rejection procedures. (These participants had fewer than 15% usable trials after artifact rejection.) Thus, statistical analyses were conducted on the processed, filtered data for the remaining 32 participants.

## RESULTS

We first report the referential ambiguity effect as indicated by the Nref in the group as a whole and then report individual differences in the Nref effect as a function of S-3 EEG activity and performance on vocabulary, WM span, and AX-CPT tasks. EEG measures and individual difference measures that significantly correlated with the Nref effect were entered into a multiple regression model to assess the unique contribution of each in predicting the size of the Nref effect.

**Figure 1.** ERPs to the critical anaphor in S-4, comparing ambiguous (solid red line) with unambiguous (dotted blue line) anaphors. Negative is plotted up.



**Table 1.** Nref Effect: Ambiguous vs. Unambiguous Anaphor in S-4

<i>Midline Column</i>								
	<i>Ambiguity</i>		<i>Ambiguity × Electrode</i>					
	<i>F(1, 31)</i>	<i>p</i>	<i>F(4, 124)</i>	<i>p</i>				
Nref: 400–800 msec	<b>6.8</b>	*	1.2	<i>ns</i>				

<i>Medial Columns</i>								
	<i>Ambiguity</i>		<i>Ambiguity × Anteriority</i>		<i>Ambiguity × Hemisphere</i>		<i>Ambiguity × Anteriority × Hemisphere</i>	
	<i>F(1, 31)</i>	<i>p</i>	<i>F(2, 62)</i>	<i>p</i>	<i>F(1, 31)</i>	<i>p</i>	<i>F(2, 62)</i>	<i>p</i>
Nref: 400–800 msec	<b>10.4</b>	**	<1	<i>ns</i>	<1	<i>ns</i>	<b>4.6</b>	*

<i>Lateral Columns</i>								
	<i>Ambiguity</i>		<i>Ambiguity × Anteriority</i>		<i>Ambiguity × Hemisphere</i>		<i>Ambiguity × Anteriority × Hemisphere</i>	
	<i>F(1, 31)</i>	<i>p</i>	<i>F(3, 93)</i>	<i>p</i>	<i>F(1, 31)</i>	<i>p</i>	<i>F(3, 93)</i>	<i>p</i>
Nref: 400–800 msec	<b>8.7</b>	**	<1	<i>ns</i>	<1	<i>ns</i>	<b>5.6</b>	**

Results for the 400- to 800-msec time window at all electrode columns tested are shown. Statistically significant values are in **bold**.

\* $p < .05$ .

\*\* $p < .01$ .

\*\*\* $p < .001$ .

### The Nref Effect

ERPs to the anaphors in S-4 (ambiguous vs. unambiguous) are displayed in Figure 1. Separate repeated-measures ANOVAs were conducted for anaphors over midline (AFz, Fz, Cz, Pz, and POz), medial (FC1, FC2, C3, C4, CP1, and CP2), and lateral (F3, F4, FC5, FC6, CP5, CP6, P3, and P4) electrode columns. Referential ambiguity was a within-participant factor. The analysis also included topographic distribution variables: For the midline column, this was electrode site (five levels), and for the medial and lateral analyses, these were hemisphere (left and right) and anteriority (medial: fronto-central, centro-parietal, parietal; lateral: frontal, fronto-central, centro-parietal, parietal). A Greenhouse–Geisser correction was used for all  $F$  tests with more than 1  $df$  in the numerator. The repeated-measures ANOVAs were conducted on the mean amplitude in the 400- to 800-msec time window, selected to be consistent with the Nref window used in previous studies (Nieuwland & Van Berkum, 2006, 2008). Results are summarized below and presented in full in Table 1.

There was a significant main effect of Referential ambiguity for all electrode columns (midline, medial, and lateral), such that ambiguous anaphors elicited a negative deflection compared with unambiguous ones. In the medial and lateral columns, this variable interacted with Anteriority and Hemisphere, reflecting the left-frontal/central maximum of the effect, as seen in Figure 1. As can also be seen in Figure 1, the effect was long lasting, lingering through the end of the epoch.

### Individual Differences in the Nref

Individual Nref effect estimates were calculated by subtracting the mean amplitude in the unambiguous condition from the ambiguous condition in the 400- to 800-msec time window. Consistent with previous work on individual differences in the Nref effect (Nieuwland & Van Berkum, 2006), a four-electrode left anterior cluster was selected, for which the Nref was maximal (F3, FC1, FC5, and C3). The average Nref effect at this cluster was used in the following analyses. The average vocabulary score was 66.5 (range = 49–79). The average listening span score was 63.9 (range = 37–90), and the average operation span score was 44.69 (range = 11–75). For the AX-CPT,  $d'$  was computed as a measure of context processing (Cohen et al., 1999), using accuracy on AX trials, controlling for errors on BX trials. This provides a measure of the ability to maintain context (having seen an A) to correctly respond to the target (X), taking into account false alarms to BX trials [ $d'$  context =  $z(\text{AX accuracy}) - z(\text{BX error})$ ]. Simple correlations were calculated among the Nref, vocabulary, WM capacity (listening span and operation span), and AX-CPT performance. Vocabulary and operation span were significantly correlated with the Nref effect ( $r = -.488, p < .01$ ;  $r = -.416, p < .05$ , respectively), such that higher scores on both measures were associated with larger Nref effects. Vocabulary was positively correlated with listening span ( $r = .358, p < .05$ ) as well as operation span ( $r = .456, p < .01$ ). Operation span was positively

**Table 2.** Simple Correlations for S-4 Nref Effect and Behavioral Measures of Vocabulary (Nelson–Denny), WM (Listening Span Total Score), and Cognitive Control (AX-CPT  $d'$  Context)

	<i>Nref Effect</i>	<i>Vocabulary</i>	<i>Listening Span</i>	<i>Operation Span</i>	<i>d' Context</i>
Nref effect					
Vocabulary	-.488***				
Listening span	-.119	.358**			
Operation span	-.416**	.456***	.237		
$d'$ context	-.282	.336*	.362	.391**	

Note that larger Nref effects are reflected by more negative values.

\* $p < .06$ .

\*\* $p < .05$ .

\*\*\* $p < .01$ .

\*\*\*\* $p < .001$ .

correlated with performance on the AX-CPT ( $r = .391, p < .05$ ). Results are summarized in Table 2.

### EEG Activity at S-3 and the Nref

An S-3 ambiguity difference (ambiguous – unambiguous) was calculated for each frequency band of interest (alpha and theta) and electrode cluster. Simple correlations were calculated among them and the Nref effect.<sup>2</sup> Correlations among electrode clusters were expected because of volume conduction across electrodes in scalp-recorded EEG; of interest here were correlations with the Nref effect. The S-3 theta difference was significantly correlated with the S-4 Nref effect at the central electrode cluster ( $r = -.392, p < .05$ ), and the S-4 alpha difference was significantly correlated with the S-4 Nref effect at all clusters but the posterior right ( $r_s = .412-.459, p_s < .05$ ). On the basis of these results, electrode clusters with significant correlations to the Nref effect were combined into a composite cluster. Thus, a central theta cluster and a composite alpha cluster (all clusters with the exception of the posterior right one) were computed and used in subsequent analyses.

### Control Analysis: Nref Effect and EEG Activity in Sentence 2

To assess whether the effect described above was specific to theta and alpha activity during S-3, an analogous anal-

ysis was conducted using the same ambiguity difference measure but calculated during an earlier point in the stories. Namely, simple correlations for the ambiguity difference during S-2 (before the critical entities being introduced) and the Nref effect were computed. There were no significant correlations between the S-4 Nref effect and any S-2 difference measure. Furthermore, the composite alpha and theta difference measures that were used to assess the relation of processing at S-3 to the Nref effect at S-4 were also computed for S-2. These were also uncorrelated with the Nref effect ( $p_s > .45$ ), and the ambiguity difference measures at S-2 and S-3 were uncorrelated with each other ( $p_s > .72$ ).

### EEG Activity and Individual Difference Measures

Simple correlations were calculated among the alpha and theta measures at S-3 and the behavioral measures—vocabulary, WM span (listening span and operation span), and performance on the AX-CPT. Only operation span was significantly correlated with the theta difference measure ( $r = .644, p < .001$ ), such that larger values on measure (ambiguous > unambiguous at S-3) were associated with better performance on operation span. No individual difference measures were significantly correlated with the S-3 alpha difference measure. Results are summarized in Table 3. The relation between the S-3 theta difference and operation span is depicted in Figure 2.

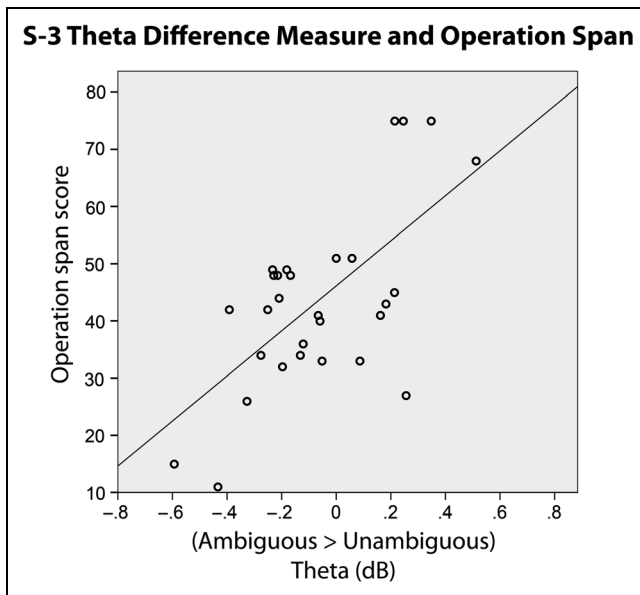
**Table 3.** Simple Correlations for the S-3 Theta and Alpha Difference Measures and Behavioral Individual-difference Measures

	<i>Vocabulary</i>	<i>Listening Span</i>	<i>Operation Span</i>	<i>d' Context</i>	<i>S-3 Alpha</i>	<i>S-3 Theta</i>
<i>S-3 EEG Measures</i>						
Alpha difference	-.114	.002	-.087	-.135		
Theta difference	.168	.225	.644****	.068	-.282	

Note that larger Nref effects are reflected by more negative values.

\*\*\*\* $p < .001$ .





**Figure 2.** Single regression correlation between the S-3 theta difference measure and operation span. On the y axis is operation span; on the x axis is the ambiguous > unambiguous S-3 difference in the theta frequency band.

### Nref Effect: Multiple Regression Analysis

We conducted a multiple regression analysis to determine which of the EEG and individual difference measures were unique in predicting the size of an individual's Nref effect (an index of sensitivity to referential ambiguity). We included the central theta difference and the composite alpha difference at S-3, vocabulary, and operation span. Vocabulary and operation span were positively correlated ( $r = .456, p < .01$ ), as were the S-3 central theta difference and operation span ( $r = .644, p < .001$ ). The predictors were otherwise uncorrelated ( $ps > .2$ ). Results of this analysis are summarized in Table 4.

**Table 4.** Multiple Regression Analysis with Four Predictors of S-4 Nref Effect (Central Theta S-3 Difference, Composite Alpha S-3 Difference, Vocabulary, and Operation Span)

<i>DV = S-4 Nref Effect</i>			
<i>Predictor</i>	<i>b</i>	$\beta$	<i>p</i>
Constant	3.436		
S-3 central theta	-0.908	-0.166	.384
S-3 composite alpha	1.451	0.458	.004**
Vocabulary	-0.055	-0.329	.041*
Operation span	-0.012	-0.131	.512

Unstandardized (*b*) and standardized ( $\beta$ ) partial coefficients and probability levels (*p*) are shown.

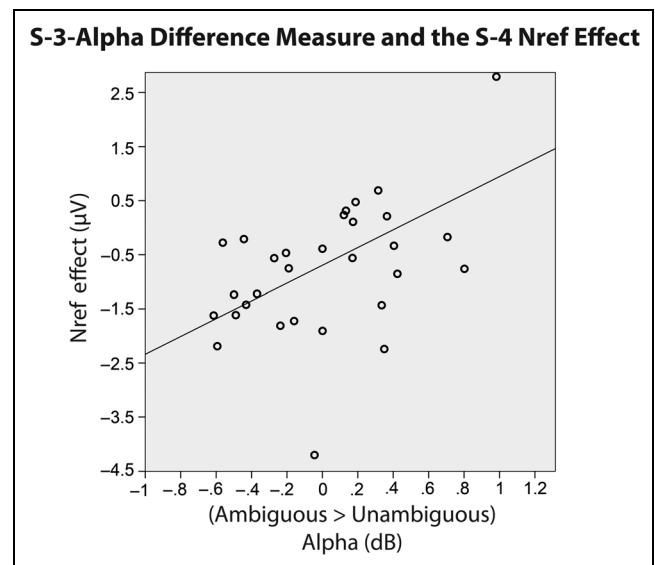
\* $p < .05$ .

\*\* $p < .01$ .

The predictors accounted for 53.9% of the variance in the S-4 Nref effect ( $R^2 = .539, F(4, 29) = 7.3, p < .001$ ). Only the S-3 alpha difference and vocabulary were significant independent predictors of the Nref effect, such that larger values on the S-3 alpha difference measure (ambiguous > unambiguous at S-3) predicted smaller Nref effects at S-4. In contrast, higher vocabulary scores predicted larger Nref effects. The relation between the alpha difference and the Nref is depicted in Figures 3 and 4. The relation between vocabulary and the Nref effect is depicted in Figure 5.

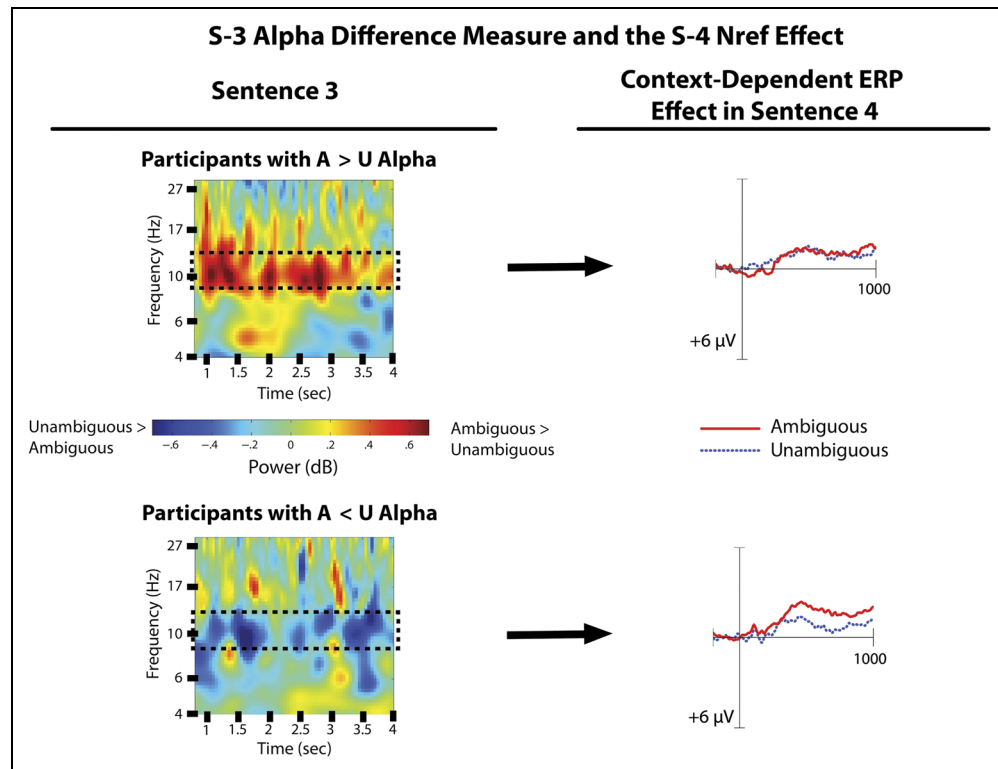
### DISCUSSION

The goal of this study was to examine the roles of attention and WM processes in the establishment of discourse reference. Specifically, we examined variation in the amplitude of the Nref effect to ambiguous versus unambiguous anaphors appearing in the final sentence of four-sentence stories as a function of EEG activity during the presentation of critical information earlier in the stories. In addition, we related variation in the Nref effect to individual differences in WM capacity (span), verbal ability (vocabulary size), and the ability to use context in performing a task (performance on the AX-CPT). In response to the critical anaphor in the fourth and final sentence of each story (S-4), we found a robust Nref effect at the group level, such that ambiguous anaphors elicited a sustained negative deflection, maximal over left fronto-central sites, compared with unambiguous anaphors. This effect reflects the processing costs associated with referential ambiguity. Our results are consistent with previous studies in which



**Figure 3.** Single regression correlation between the S-3 alpha difference measure and the S-4 Nref effect. On the y axis is the S-4 Nref effect; on the x axis is the ambiguous > unambiguous S-3 difference in the alpha frequency band. Note that larger negative values for the Nref effect indicate larger effect sizes.

**Figure 4.** Relation between S-3 ambiguous > unambiguous alpha difference and S-4 Nref effect. For display purposes only, participants were split into A > U ( $N = 16$ ) and A < U ( $N = 16$ ) groups based on the median S-3 alpha difference value ( $\sim 0$ ). On the left are time-frequency plots showing spectral power for the S-3 difference across all frequencies examined; the alpha frequency band (9–12 Hz) is highlighted within the black boxes for the A > U group (top) and the A < U group (bottom). On the right are the corresponding S-4 ERP results from a representative electrode in the left-anterior electrode cluster (FC5), comparing ambiguous (solid red line) with unambiguous (dotted blue line) S-4 anaphors for the A > U group (top) and the A < U group (bottom).

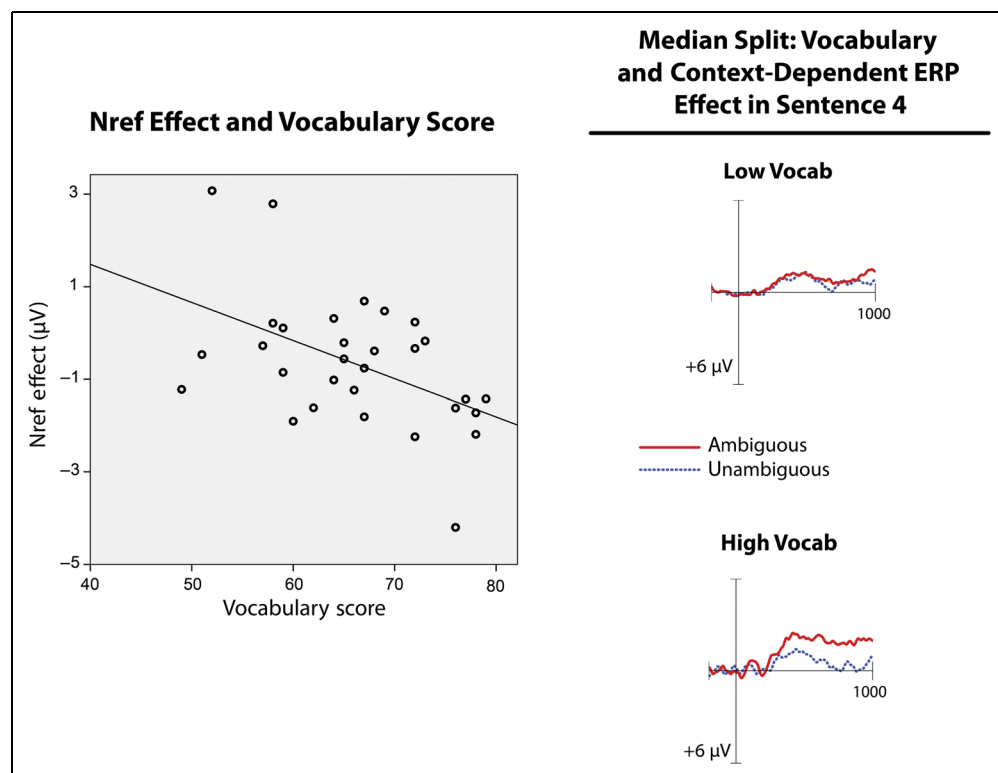


an Nref effect was found in response to manipulations of referential context (Nieuwland et al., 2007; Nieuwland & Van Berkum, 2006; Van Berkum et al., 1999, 2003).

The results of our multiple regression analysis demonstrated that the Nref effect varied across individuals and was significantly predicted by individual differences in

vocabulary and EEG power in the alpha frequency band. Individuals with large vocabularies were more sensitive to the referential ambiguity. This likely reflects the relation between vocabulary size and comprehension skill. Individuals who are skilled in the comprehension of language tend to have large vocabularies and tend to be more

**Figure 5.** (Left) Single regression correlation between vocabulary score and the S-4 Nref effect. On the y axis is the S-4 Nref effect; on the x axis is the vocabulary score. Note that larger negative values for the Nref effect indicate larger effect sizes. (Right) For display purposes only, participants were split into high vocabulary ( $N = 14$ ) and low vocabulary ( $N = 14$ ) groups based on the median score (scores at the median were discarded). S-4 ERP results from the left-anterior electrode cluster, comparing ambiguous (solid red line) with unambiguous (dotted blue line) referents, are shown for the high vocab group (top) and the low vocab group (bottom).



sensitive to linguistic ambiguity of all types (Long et al., 2006). The relation between alpha power and the Nref was such that individuals who were sensitive to the ambiguity of the S-4 anaphor tended to show increased alpha power at S-3 in the ambiguous condition compared with the unambiguous condition. Importantly, the relation between alpha and the Nref was specific to alpha activity that occurred when critical information was introduced in S-3. Alpha power at S-2 had no relation to the size of the Nref effect.

With respect to individual variation in EEG activity, we found that higher scores on a WM span measure (operation span) were predictive of increased theta power in the ambiguous condition (two oaks) compared with the unambiguous condition (an oak and an elm) at S-3.

### **Alpha and Attentional Engagement during Discourse Processing**

As noted in the Introduction, increases in alpha power have been linked to decreases in attentional engagement to external stimuli (Roux & Uhlhaas, 2014; Meyer et al., 2013; Bengson et al., 2012; Jensen & Mazaheri, 2010; Thut et al., 2006; Jensen et al., 2002), but few studies have investigated the role of alpha in language processing, particularly at the discourse level. In the current study, we investigated whether alpha power, measured while participants listened to soon-to-be critical information, was predictive of the subsequent ability to use that information to process an anaphor. If alpha power indexes attentional engagement to an external task (story listening, in this study), then individuals who exhibit relatively high alpha power during the presentation of important versus unimportant information should be less sensitive to a manipulation that depends on having encoded that information.

This was the pattern of results that we observed in the current study: Individuals with larger differences in alpha at S-3 (ambiguous > unambiguous) were less sensitive to the ambiguity of a subsequent anaphor. It is important to note that individuals with larger compared with smaller values on the alpha difference measure did not necessarily exhibit increases in alpha power overall. Rather, individuals with high values tended to have increases in alpha power that occurred at S-3 in the ambiguous condition. Critical information was presented at S-3; thus, attentional disengagement during that time had consequences for processing the anaphor in the following sentence. This result is consistent with studies of visual attention and WM in which increases in alpha power are found when individuals ignore input from the external environment (which, in those tasks, serves to selectively focus attention on task-relevant information).

As discussed in the Introduction, in many tasks, inhibition of external input promotes optimal performance, as when participants ignore one location in visual space to respond to a target at another location (e.g., Bengson

et al., 2012; Thut et al., 2006) or when individuals focus on to-be-remembered information during the maintenance period in a WM task (e.g., Jensen et al., 2002). Although it can be advantageous to ignore external input in a task, such as during the delay period in a WM task, suppression of external input while listening to a story is likely to be problematic even when WM demands are high. If listeners ignore input to focus on internal operations and representations during language comprehension, then critical information can be missed. In our task, inhibition of external input during the presentation of critical information at S-3 would result in failure to recognize the referential ambiguity in S-4.

Thus, the correlation between alpha power and the Nref in our study is consistent with the hypothesis that increases in alpha power reflect attentional disengagement from the external task (story listening). Some researchers, however, have argued that alpha power increases are associated with task-relevant semantic retrieval or WM processes. This account of alpha is not consistent with the pattern of results in this study, because alpha increases were associated with impairments of comprehension. If alpha increased during S-3 in the ambiguous condition because listeners were engaged in processes that were relevant for establishing and maintaining representations of the critical information, then alpha and the Nref should have been positively related: Larger values on the alpha difference measure should have been associated with greater sensitivity to referential ambiguity. This is because the ambiguity of the anaphor should be glaringly obvious in light of an accurate and strong representation of two possible antecedents in the discourse (oak/oak). Thus, we interpret our results as indicating that the Nref in our study was associated with the quality of the contextual representations of the ambiguous condition. This interpretation is consistent with our finding that vocabulary and the Nref were correlated. Individuals with higher vocabulary scores (an important predictor of verbal skill) showed larger Nref effects than individuals with lower scores, suggesting that comprehension ability is associated with greater sensitivity to referential ambiguity. Similarly, Nieuwland and Van Berkum (2006) found that individuals with higher scores on a WM span task showed larger Nref effects of referential ambiguity. Both of these results are consistent with the idea that enhanced context representations (or access to context representations) lead to larger Nref effects. These findings are inconsistent with the interpretation that increased alpha in this study reflected enhanced processing of task-relevant input.

We therefore suggest that increases in alpha power in this study reflected reductions in attention to the processing of incoming words. An important question is whether attentional disengagement was a characteristic of some individuals in the study or was indicative of momentary lapses of attention that, for some individuals, happened to occur at inopportune times, that is, during

S-3 in the ambiguous condition. We favor the latter possibility for two reasons. First, we used a within-participant measure to predict individual differences in sensitivity to referential ambiguity. This means, as described above, that it was not the overall estimate of alpha power for a given participant that predicted the Nref effect but rather the individual's alpha estimate on the ambiguous condition relative to the unambiguous condition. Thus, it was the pattern of fluctuations in alpha (used here as a proxy for attentional engagement to external input) within an individual over the course of the experiment that had predictive power, rather than the individual's overall level of attentiveness. Second, we examined the relation between the same alpha difference measure at S-2 and the size of the Nref. If some individuals paid attention throughout the story, whereas other individuals did not, then the Nref would likely be related to attentiveness at both S-2 and S-3. Our results showed, however, that the relation between alpha and the Nref was specific to attention to input at S-3 when critical information was presented about possible antecedents of the subsequent pronoun. Thus, we suggest that alpha differences at S-3 captured variability associated with fluctuations in attention to the language input within an individual, rather than his or her overall level of attentional (dis)engagement during story comprehension.

Although our results suggest that lapses in attention to external input during listening can impair comprehension, they provide no information about why the input was ignored. One possibility is that these momentary dips in attention to external input corresponded to episodes of mind wandering in which individuals shifted their attention from processing incoming words to processing task-irrelevant thoughts and concerns. Another possibility is that individuals shifted their attention from the input to task-relevant thoughts. For example, attention to incoming language input might lapse when listeners encounter difficulty in processing because they need to focus on resolving this difficulty to comprehend the story. Moreover, lapses of attention to external input can be spontaneous or deliberate. Although the terms "mind-wandering," "task-unrelated thought," and "spontaneous thought" are often used interchangeably, there are some important differences among them. For example, spontaneous thought is by definition undirected and comes unbidden to the forefront of attention, whereas it is possible to mind wander quite purposefully and about something specific (Christoff, 2012). The results of the current study show that lapses of attention to the language input occurred and that these lapses had consequences for subsequent processing. We cannot determine whether these lapses were spontaneous or deliberate, nor can it be determined whether individuals were engaged in task-related or task-unrelated thought during these intervals. Future studies will be needed in which the content of individuals' thoughts is assessed to address these issues.

An important question for future research concerns the relation between attentional lapses during language processing and the process(es) by which individuals "snap back" to attention. It may be that readers/listeners periodically monitor their comprehension and reengage when they realize that they are off-task. Alternatively, some characteristic of the input, such as a change in prosody, might serve to recapture attention.<sup>3</sup> Studies are currently underway to investigate how properties of the linguistic input may refocus attention. Individual differences may also play a role; individuals who are sensitive to contextual cues indicating the importance of a piece of information (e.g., prosody, syntactic structures that place an entity in discourse focus) may respond to these cues by refocusing attention (potentially reducing alpha oscillations). In the current study, the alpha difference at S-3 was uncorrelated with the individual differences measures (vocabulary, WM span, and AX-CPT performance), although vocabulary did account for a unique share of variance in the Nref effect. However, this does not preclude the possibility that individual differences in these measures would affect attentional engagement given different types of manipulations (e.g., difficult sentence structures) or that variability in attentional engagement is influenced by other individual differences, such as motivation or interest in the story.

### **Theta and WM during Discourse Processing**

Oscillations in the theta band have increasingly been associated with memory and WM processes (Hsieh et al., 2011; Sauseng et al., 2010; Meltzer et al., 2008; Hald et al., 2006; Bastiaansen et al., 2002a, 2002b; Jensen & Tesche, 2002; Gevins & Smith, 2000; Gevins et al., 1997). The results of the current study showed that higher scores on a WM span measure (operation span) were predictive of increased theta differences at S-3. We hypothesized that processing at S-3 in the ambiguous condition would be more demanding than processing in the unambiguous condition. This was because the entities in the ambiguous condition could be distinguished only by processing the modifying phrases (oak with the mushroom on it/oak with birds in its branches), whereas the entities in the unambiguous condition could be distinguished based on the meanings of the words used to refer to them. The finding that higher WM span individuals showed a larger power difference in theta across conditions is consistent with a recent study in which higher frontal theta power was specifically associated with WM maintenance for the order of items rather than the items themselves (Hsieh et al., 2011). In that study, frontal theta power was increased when participants were asked to track a sequence of kaleidoscope images, as compared with processing the images themselves; the effect was enhanced for participants who performed especially well on this visual sequencing WM task (Hsieh et al., 2011). Similarly, in the current experiment, higher WM span participants had increased theta power in the ambiguous condition when

processing demands were high. Individual differences in theta power under high WM load conditions have also been linked to performance on tests of cognitive ability, including WM and verbal tasks (Gevins & Smith, 2000).

Finally, the theta measure (ambiguous > unambiguous) at S-3 was significantly correlated with the size of the Nref effect, such that increases in the theta difference measure were linked to larger Nref effects. However, the results of the multiple regression analysis showed that the theta difference measure did not account for unique variance in predicting the Nref effect when other predictors were included. Thus, the correlation between theta and the Nref may be secondary to shared variance with other variables.

## Conclusions

We found that individual variation in verbal ability and attention significantly predicted sensitivity to referential ambiguity. Higher vocabulary scores predicted larger Nref effects, whereas increased alpha power while listening to critical information in a sentence that rendered a subsequent anaphor ambiguous predicted smaller Nref effects. We suggest that the alpha difference at S-3 in the current study captured the waxing and waning of attention to the stories, which had an impact on discourse-level referential processing. Listeners who tended to be inattentive when the antecedents were presented tended to be insensitive to referential ambiguity. In addition, WM span was predictive of theta power during S-3, such that individuals with higher spans showed a larger theta increase in the more demanding condition than in the less demanding one. Overall, the pattern of results provides evidence that the engagement of WM processes during language comprehension can be indexed by changes in theta-band power, whereas fluctuations in attention to incoming words can be indexed by changes in alpha-band power. These results show that the establishment of reference places demands on attention and WM processes. Connecting an anaphor to its appropriate antecedent requires, first, that attention be directed to the relevant information when it is introduced. Second, the antecedents of anaphors need to be represented in WM with sufficient detail such that they can be uniquely identified when an anaphor is processed.

Reprint requests should be sent to Megan A. Boudewyn, Department of Psychology, Center for Neuroscience, Center for Mind and Brain, UC Davis, 1 Shields Avenue, Davis, CA 95616, or via e-mail: maboudewyn@ucdavis.edu.

## Notes

1. After the critical referential expression in S-4, the stories continued and contained a syntactic manipulation that is not relevant to the current study and has been reported elsewhere (Dave, Boudewyn, & Swaab, 2013).
2. We also examined oscillatory activity in the beta frequency band based on recent research suggesting that activity in this band is associated with language processing and, in particular,

with the processes involved in syntactic analysis (Weiss & Mueller, 2012; Bastiaansen, Magyari, & Hagoort, 2010; Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008; Bastiaansen & Hagoort, 2006; Bastiaansen, Van Der Linden, Ter Keurs, Dijkstra, & Hagoort, 2005; Weiss et al., 2005; Haarmann, Cameron, & Ruchkin, 2002). We examined the S-3 ambiguity difference in the beta band (14–18 Hz) in relation to the S-4 Nref effect. The difference in the beta band was not correlated with any of our measures; thus, these results are not discussed.

3. We refer to fluctuations in attention over time that are not specifically driven by physiological states such as hunger or psychological competitors for attention such as intrusive thoughts or hallucinations. These types of attention lapses no doubt play a role in the ability to allocate attention to a task. Recent work has also shown that cigarette craving and the consumption of alcohol increase the incidence of attention lapses while reading (Sayette, Schooler, & Reichle, 2010; Sayette, Reichle, & Schooler, 2009). In the current article, we are primarily focused on lapses in attention that occur in the absence of these factors.

## REFERENCES

- Adrian, E. D., & Matthews, B. H. (1934). The Berger rhythm: Potential changes from the occipital lobes in man. *Brain*, *57*, 355–385.
- Bastiaansen, M. C., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in Brain Research*, *159*, 179–196.
- Bastiaansen, M. C., Magyari, L., & Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *Journal of Cognitive Neuroscience*, *22*, 1333–1347.
- Bastiaansen, M. C., Oostenveld, R., Jensen, O., & Hagoort, P. (2008). I see what you mean: Theta power increases are involved in the retrieval of lexical semantic information. *Brain and Language*, *106*, 15–28.
- Bastiaansen, M. C., Van Berkum, J. J. A., & Hagoort, P. (2002a). Event-related theta power increases in the human EEG during online sentence processing. *Neuroscience Letters*, *323*, 13–16.
- Bastiaansen, M. C., Van Berkum, J. J. A., & Hagoort, P. (2002b). Syntactic processing modulates the  $\theta$  rhythm of the human EEG. *Neuroimage*, *17*, 1479–1492.
- Bastiaansen, M. C., Van Der Linden, M., Ter Keurs, M., Dijkstra, T., & Hagoort, P. (2005). Theta responses are involved in lexical—Semantic retrieval during language processing. *Journal of Cognitive Neuroscience*, *17*, 530–541.
- Bengson, J. J., Mangun, G. R., & Mazaheri, A. (2012). The neural markers of an imminent failure of response inhibition. *Neuroimage*, *59*, 1534–1539.
- Blozis, S. A., & Traxler, M. J. (2007). Analyzing individual differences in sentence processing performance using multilevel models. *Behavior Research Methods*, *39*, 31–38.
- Christoff, K. (2012). Undirected thought: Neural determinants and correlates. *Brain Research*, *1428*, 51–59.
- Cohen, J. D., Barch, D. M., Carter, C., & Servan-Schreiber, D. (1999). Context-processing deficits in schizophrenia: Converging evidence from three theoretically motivated cognitive tasks. *Journal of Abnormal Psychology*, *108*, 120.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.
- Dave, S., Boudewyn, M. A., & Swaab, T. Y. (2013). *Individual differences in the effects of referential context on syntactic ambiguity resolution*. Paper presented at the Meeting of the Cognitive Neuroscience Society, Boston.

- Dockree, P. M., Kelly, S. P., Foxe, J. J., Reilly, R. B., & Robertson, I. H. (2007). Optimal sustained attention is linked to the spectral content of background EEG activity: Greater ongoing tonic alpha (~10 Hz) power supports successful phasic goal activation. *European Journal of Neuroscience*, *25*, 900–907.
- Dockree, P. M., Kelly, S. P., Roche, R. A., Hogan, M. J., Reilly, R. B., & Robertson, I. H. (2004). Behavioural and physiological impairments of sustained attention after traumatic brain injury. *Cognitive Brain Research*, *20*, 403–414.
- Franklin, M. S., Broadway, J. M., Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2013). Window to the wandering mind: Pupillometry of spontaneous thought while reading. *The Quarterly Journal of Experimental Psychology*, *66*, 2289–2294.
- Franklin, M. S., Smallwood, J., & Schooler, J. W. (2011). Catching the mind in flight: Using behavioral indices to detect mindless reading in real time. *Psychonomic Bulletin & Review*, *18*, 992–997.
- Gevins, A., & Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex*, *10*, 829–839.
- Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: Effects of task difficulty, type of processing, and practice. *Cerebral Cortex*, *7*, 374–385.
- Haarmann, H. J., Cameron, K. A., & Ruchkin, D. S. (2002). Neural synchronization mediates on-line sentence processing: EEG coherence evidence from filler-gap constructions. *Psychophysiology*, *39*, 820–825.
- Hald, L. A., Bastiaansen, M. C., & Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. *Brain and Language*, *96*, 90–105.
- Hamilton, S. T., Freed, E. M., & Long, D. L. (2013). Modeling reader and text interactions during narrative comprehension: A test of the lexical quality hypothesis. *Discourse Processes*, *50*, 139–163.
- Hsieh, L.-T., Ekstrom, A. D., & Ranganath, C. (2011). Neural oscillations associated with item and temporal order maintenance in working memory. *Journal of Neuroscience*, *31*, 10803–10810.
- Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the alpha band (9–12 Hz) increase with memory load during retention in a short-term memory task. *Cerebral Cortex*, *12*, 877–882.
- Jensen, O., & Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: Gating by inhibition. *Frontiers in Human Neuroscience*, *4*, 186.
- Jensen, O., & Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *European Journal of Neuroscience*, *15*, 1395–1399.
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, *29*, 169–195.
- Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in Cognitive Sciences*, *16*, 606–617.
- Long, D. L., Johns, C., & Morris, P. (2006). Comprehension ability in mature readers. *Handbook of Psycholinguistics*, *2*, 801–833.
- Long, D. L., Prat, C., Johns, C., Morris, P., & Jonathan, E. (2008). The importance of knowledge in vivid text memory: An individual-differences investigation of recollection and familiarity. *Psychonomic Bulletin & Review*, *15*, 604–609.
- Mazaheri, A., Coffey-Corina, S., Mangun, G. R., Bekker, E. M., Berry, A. S., & Corbett, B. A. (2010). Functional disconnection of frontal cortex and visual cortex in attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *67*, 617–623.
- Meltzer, J. A., Zaveri, H. P., Goncharova, I. I., Distasio, M. M., Papademetris, X., Spencer, S. S., et al. (2008). Effects of working memory load on oscillatory power in human intracranial EEG. *Cerebral Cortex*, *18*, 1843–1855.
- Meyer, L., Obleser, J., & Friederici, A. D. (2013). Left parietal alpha enhancement during working memory-intensive sentence processing. *Cortex*, *49*, 711–721.
- Nieuwland, M. S., Otten, M., & Van Berkum, J. J. (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., & Van Berkum, J. J. A. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research*, *1118*, 155–167.
- Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and Language*, *106*, 119–131.
- O’Connell, R. G., Dockree, P. M., Bellgrove, M. A., Turin, A., Ward, S., Foxe, J. J., et al. (2009). Two types of action error: Electrophysiological evidence for separable inhibitory and sustained attention neural mechanisms producing error on go/no-go tasks. *Journal of Cognitive Neuroscience*, *21*, 93–104.
- O’Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: Electrophysiological signals predict errors up to 20 s before they occur. *Journal of Neuroscience*, *29*, 8604–8611.
- Pfurtscheller, G., Stancak, A., Jr., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band—An electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology*, *24*, 39–46.
- Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science*, *21*, 1300–1310.
- Roche, R. A., Dockree, P. M., Garavan, H., Foxe, J. J., Robertson, I. H., & O’Mara, S. M. (2004). EEG alpha power changes reflect response inhibition deficits after traumatic brain injury (TBI) in humans. *Neuroscience Letters*, *362*, 1–5.
- Röhm, D., Klimesch, W., Haider, H., & Doppelmayr, M. (2001). The role of theta and alpha oscillations for language comprehension in the human electroencephalogram. *Neuroscience Letters*, *310*, 137–140.
- Roux, F., & Uhlhaas, P. J. (2014). Working memory and neural oscillations: Alpha-gamma versus theta-gamma codes for distinct WM information? *Trends in Cognitive Sciences*, *18*, 16–25.
- Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience & Biobehavioral Reviews*, *34*, 1015–1022.
- Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R., & Hanslmayr, S. (2005). EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Human Brain Mapping*, *26*, 148–155.
- Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, *57*, 97–103.
- Sayette, M. A., Reichle, E. D., & Schooler, J. W. (2009). Lost in the sauce: The effects of alcohol on mind wandering. *Psychological Science*, *20*, 747–752.
- Sayette, M. A., Schooler, J. W., & Reichle, E. D. (2010). Out for a smoke: The impact of cigarette craving on zoning out during reading. *Psychological Science*, *21*, 26–30.

- Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind: Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review*, *14*, 230–236.
- Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition*, *36*, 1144–1150.
- Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006).  $\alpha$ -Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *Journal of Neuroscience*, *26*, 9494–9502.
- Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory*, *17*, 635–654.
- Van Berkum, J. J. A. (2008). Understanding sentences in context what brain waves can tell us. *Current Directions in Psychological Science*, *17*, 376–380.
- Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*, 147–182.
- Van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event related brain potentials reflect discourse referential ambiguity in spoken language comprehension. *Psychophysiology*, *40*, 235–248.
- Van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, *1146*, 158–171.
- Van Winsum, W., Sergeant, J., & Geuze, R. (1984). The functional significance of event-related desynchronization of alpha rhythm in attentional and activating tasks. *Electroencephalography and Clinical Neurophysiology*, *58*, 519–524.
- Waters, G. S., & Caplan, D. (2003). The reliability and stability of verbal working memory measures. *Behavior Research Methods, Instruments, & Computers*, *35*, 550–564.
- Weiss, S., & Mueller, H. M. (2012). “Too many betas do not spoil the broth”: The role of beta brain oscillations in language processing. *Frontiers in Psychology*, *3*, 201.
- Weiss, S., Mueller, H. M., Schack, B., King, J. W., Kutas, M., & Rappelsberger, P. (2005). Increased neuronal communication accompanying sentence comprehension. *International Journal of Psychophysiology*, *57*, 129–141.