Investigating the Time Course of Spoken Word Recognition: Electrophysiological Evidence for the Influences of Phonological Similarity

Amy S. Desroches[1], Randy Lynn Newman[2], and Marc F. Joanisse[1]

Abstract

■ Behavioral and modeling evidence suggests that words compete for recognition during auditory word identification, and that phonological similarity is a driving factor in this competition. The present study used event-related potentials (ERPs) to examine the temporal dynamics of different types of phonological competition (i.e., cohort and rhyme). ERPs were recorded during a novel picture–word matching task, where a target picture was followed by an auditory word that either matched the target (CONE–cone), or mismatched in one of three ways: rhyme (CONE–bone), cohort (CONE–comb), and unrelated (CONE–fox). Rhymes and cohorts differentially modulated two distinct ERP components, the phonological mismatch negativity and the N400, revealing the influences of prelexical and lexical processing components in speech recognition. Cohort mismatches resulted in late increased negativity in the N400, reflecting disambiguation of the later point of miscue and the combined influences of top–down expectations and misleading bottom–up phonological information on processing. In contrast, we observed a reduction in the N400 for rhyme mismatches, reflecting lexical activation of rhyme competitors. Moreover, the observed rhyme effects suggest that there is an interaction between phoneme-level and lexical-level information in the recognition of spoken words. The results support the theory that both levels of information are engaged in parallel during auditory word recognition in a way that permits both bottom–up and top–down competition effects. ■

INTRODUCTION

In understanding spoken language, listeners need to both perceive incoming auditory information and access a semantic representation of that input. Although speech is understood quite rapidly and effortlessly, the cognitive processing involved in spoken word recognition is not trivial. In order to recognize what is being said, acoustic information must be translated into a phonological code, segmented into discrete words, and integrated with both the immediate context and prior knowledge such as word familiarity and contextual cues. Consistent with this, studies have revealed a range of factors that influence auditory word recognition (Norris, McQueen, & Cutler, 2000; Luce & Pisoni, 1998; Frauenfelder & Tyler, 1987; McClelland & Elman, 1986). These can be coarsely divided into two categories: those revealing the influence of "prelexical" cues related to acoustic and phonological features, and those indexing the role of "lexical" knowledge that denotes lexical-level influences such as frequency. There is significant ongoing discussion about exactly how auditory words are recognized, focusing on how and when these different types of information are accessed during the time course of processing, and furthermore, the extent to which these interact.

Evidence suggests that spoken words are processed as speech unfolds, and that phonologically similar items compete for recognition during spoken word identification (Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Norris et al., 2000; Vitevitch & Luce, 1999; Allopenna, Magnuson, & Tanenhaus, 1998; Marslen-Wilson & Zwitserlood, 1989; McClelland & Elman, 1986; Marslen-Wilson & Tyler, 1980). Word-initial phonological overlap results in cohort interference, with words sharing the same initial sounds (e.g., *cap*, *cat*, *cab*, *catch*, and *captain*, termed "cohorts") competing for recognition (Marslen-Wilson & Zwitserlood, 1989). Further influences of phonological similarity are also observed, with other so-called neighbors showing competition effects (i.e., words that differ from *cap* by only one phoneme, such as *cop*, *cape*, and *clap*, which also includes rhymes, *map*, *tap*, *zap*; Luce & Pisoni, 1998). Studies that have examined cohort effects and/or global neighborhood effects have suggested that interference makes words with many neighbors more difficult to recognize than words with few neighbors (Luce & Pisoni, 1998). In addition to the size of the competitor set, other factors such as the relative frequency of these neighbors can influence the time course of spoken word recognition (Luce & Pisoni, 1998). These data suggest that phonological information is integrated continuously during spoken word recognition, leading to competition among phonologically related words.

[1]The University of Western Ontario, [2]Acadia University

A number of models have been proposed to explain the process of spoken word recognition and to account for phonological competition. The Cohort model (Marslen-Wilson & Tyler, 1980) suggests that competing candidates become activated in spoken word recognition. The competitor set is increasingly constrained as a word unfolds, and recognition occurs when only one candidate remains. In this model, competition effects are bottom–up, occurring only among words that overlap from the initial phonemes onwards (''cohorts''). Other models such as Shortlist/Merge (Norris et al., 2000; Norris, 1994), the Neighborhood Activation Model (NAM; Luce & Pisoni, 1998), and TRACE (McClelland & Elman, 1986) permit a broader competitor set, allowing for phonological competition amongst noncohorts, such as rhymes and other neighbors. Each of these account for competition differently, with some being arguably better able to explain certain effects than others. For instance, although NAM allows for different types of phonological competition, it does not take into account the temporal nature of speech; instead, similarity is computed mathematically as the overall perceptual and phonological difference among competitors.

Continuous mapping models of spoken word recognition, such as TRACE (McClelland & Elman, 1986) and also Shortlist/Merge (Norris et al., 2000; Norris, 1994), incorporate components of both these approaches, proposing that competition occurs via lateral inhibition between lexical candidates. These models assume that recognizing a word involves activating its unique word-form representation based on acoustic–phonetic inputs. Although these inputs are provided to the system in a serial fashion, they allow for competition effects that occur due to similarity at any point in a word. However, despite some similarities, the underlying architecture of TRACE and Shortlist/Merge are quite different, especially in the way in which they account for competition. Shortlist/Merge is feedforward, having only bottom–up connections between the phoneme and lexical levels of representation. Although it provides an account for different types of competition, it is proposed that these arise from the influence of lexical knowledge on phonological processing at the decision stage. On the other hand, TRACE emphasizes the temporal and dynamic nature of speech, and in doing so is able to account for different types of lexical competition. Under this theory, reciprocal connections exist between lexical and sublexical layers permitting both bottom–up and top–down effects during word recognition.

Because a primary way in which these models differ is in how they account for phonological competition, competition effects may provide key insight into understanding the process of spoken word recognition. Although cohort effects have been consistently found in the behavioral priming literature, rhyme effects have been more elusive (e.g., Marslen-Wilson, Moss, & van Halen, 1996; Connine, Blasko, & Titone, 1993; Marslen-Wilson & Zwitserlood, 1989). Only small effects have been observed, generally using cross-modal priming, where nonwords have been shown to prime rhyme targets (e.g., *pomato**–TOMATO). Although this illustrates rhyme priming, these findings are argued to underestimate lexical competition because there is no lexical entry for nonwords (see Allopenna et al., 1998).

The difficulties in isolating rhyme competition effects, coupled with the finding of robust cohort effects, might support the theory that phonological interference occurs in a linear fashion as predicted by the Cohort model. However, compelling evidence from an eyetracking methodology called the *visual-world paradigm* has provided finer-grained evidence in this regard. Allopenna et al. (1998) studied phonological competition during spoken language processing in real time by monitoring eye gaze to a visual display during an auditory word recognition task. Participants tend to fixate pictures that depict a word as it is being heard. In addition, however, a significant proportion of fixations are also directed at phonological competitors in the display. For example, when hearing *candle*, participants looked at phonological competitors such as *candy* (a cohort competitor) and *sandal* (a rhyme competitor) more frequently than unrelated distractors like *beaker*. This finding seems to provide direct evidence for phonological competition in the time course of spoken word recognition. Looks to rhyme competitors tended to occur later than looks to cohort competitors, reflecting the fact that rhyme overlap occurs at a later point in a word. The success of this paradigm in revealing rhyme competition appears to be due to the measurement of speech processing as it unfolds. In contrast, studies using reaction time measures capture the endpoint of processing, and thus, might not be sensitive enough to reveal more subtle similarity effects.

By monitoring spoken language processing in real time, eyetracking has been successful in revealing that both cohorts and rhymes compete during recognition. In addition, these effects have also been demonstrated in the absence of visually presented competitors by manipulating the neighborhood size of targets (Magnuson, Dixon, Tanenhaus, & Aslin, 2007; Magnuson, Tanenhaus, Aslin, & Dahan 2003). Taken together, these results provide support for the view that spoken word recognition is both a continuous and dynamic process. That is, eyetracking studies have observed that top–down information (i.e., visual context) interacts with bottom–up information (i.e., auditory words) during spoken word processing, in line with the predictions of the TRACE model. However, there remains ongoing debate regarding the existence of interaction between phoneme-level and lexical-level representations, and whether or not this type of top–down feedback (i.e., from word-level to phoneme-level) is required to explain these effects (see Magnuson, Strauss, & Harris, 2005; Norris et al., 2000).

Electrophysiological Measures of Spoken Word Recognition

Event-related potentials (ERPs) can provide further insights into our understanding of the mechanisms involved in spoken word recognition. First, ERPs offer a high degree of temporal precision, thus, like eye tracking, they allow us to measure spoken word processing as it unfolds. Furthermore, certain electrophysiological components have been tied to distinct aspects of processing (i.e., phoneme level/prelexical vs. lexical level) as we discuss further below. Thus, this methodology might allow us to disentangle how prelexical and lexical processing each contribute to phonological competition effects. In doing so, this investigation promises to shed light onto the debate regarding the role of interaction in the process of spoken word recognition.

The present study investigates the role that phonological similarity plays in auditory word recognition by examining the mechanisms underlying phonological competition. We take advantage of two electrophysiological correlates of speech processing that appear to dissociate lexical and prelexical effects, the N400 (Kutas & Hillyard, 1984) and the phonological mismatch negativity (PMN) (Connolly & Phillips, 1994; see also Connolly, in press, who suggests this is better named ''Phonological Mapping Negativity''). Each of these components is characterized by a divergence in the EEG wave elicited when incoming auditory information violates an expectation; critically, they do so in different ways.

The N400 is a negative-going component occurring approximately 400 msec poststimulus onset and tends to have a central–parietal distribution. It is sensitive to incongruities in semantic or lexical information in words or sentences (Connolly & Phillips, 1994; Holcomb & Neville, 1991; Kutas & Hillyard, 1984). Furthermore, this component responds similarly during spoken word, written word, and even picture identification tasks, suggesting that it reflects lexical or semantic integration in a modality-independent way.

In contrast, the PMN[1] is thought to index prelexical processing. This component is characterized by an earlier-going negativity occurring between 250 and 300 msec poststimulus onset, typically with a midline fronto-central distribution. The PMN is specifically sensitive to differences in the expected versus perceived phonological form of a word (Newman, Connolly, Service, & McIvor, 2003; D'Arcy, Connolly, & Crocker, 2000; Hagoort & Brown, 2000; Connolly & Phillips, 1994). It is only seen during spoken word tasks and not in the visual modality, supporting the assertion that it specifically reflects an auditory phoneme matching process (Connolly & Phillips, 1994; Connolly, Phillips, Stewart, & Brake, 1992).

It has been argued that the PMN and N400 components index distinct levels of processing. Connolly and Phillips (1994) demonstrated the divergence of these two components during a sentence listening task in which the terminal word was manipulated to either meet or violate participants' expectations. When the terminal word differed from the expected high-cloze frequency word (e.g., ''the pizza was too hot to *sing*''), both a PMN and an increased N400 component were observed. Moreover, a semantically plausible word that was, nevertheless, unexpected (e.g., ''the pig wallowed in the *pen*,'' which is semantically plausible but has a much lower cloze frequency than the expected word ''mud'') yielded only a PMN and not an N400. Finally, a semantically incongruous word that was phonologically similar to the expected word yielded only an increased N400 and no PMN (e.g., ''the gambler had a streak of bad *luggage*,'' in which the semantically mismatching word is phonologically similar to the expected word ''luck''). Given this, the authors hypothesized that the PMN component specifically reflects phonological mapping because it is a negativity elicited only when initial phonological information mismatches from expectation. It is dissociable from the N400, which is thought to be related to accessing lexical and semantic information, as it can be observed even when the semantic information is in accordance with expectations (D'Arcy et al., 2000; Connolly & Phillips, 1994).

The PMN appears to be specifically sensitive to prelexical phonological information: for example, it is not sensitive to lexical status such that it occurs for phonological mismatches occurring in either word or nonword stimuli (Newman et al., 2003). In addition to being observed in sentence listening and priming paradigms, the PMN has been demonstrated using a phoneme deletion task (judging the auditory sentence ''*clap without [k] is lap*''), where the expectation is generated based on the product of phonological manipulations and judgments (Newman et al., 2003). These findings indicate that the PMN reflects the influence of top–down expectancies on bottom–up phoneme processing.

The N400 shows somewhat different characteristics. Although it has traditionally been thought of as reflecting semantic processing, a number of studies have shown that it is sensitive to a number of lexical factors including word frequency, morphological structure, and phonology (van den Brink, Brown, & Hagoort, 2001; Münte, Say, Clahsen, Schiltz, & Kutas, 1999; Van Petten, Coulson, Rubin, Plante, & Parks, 1999; Radeau, Besson, Fonteneau, & Castro, 1998; Praamstra, Meyer, & Levelt, 1994; Holcomb & Neville, 1991; Van Petten & Kutas, 1990; Kutas & Hillyard, 1984). The N400 amplitude tends to increase for semantically incongruous words, and is reduced in cases of semantic, morphological, and phonological priming. In studies using unimodal auditory priming, reductions in the N400 are seen to word targets primed by either cohorts (sometimes called alliterative priming) or rhymes (Dumay et al., 2001; Radeau et al., 1998; Praamstra et al., 1994). Importantly, a reduced N400 due to phonological priming appears to reflect the ease with which a word is retrieved, with advantages

for processing similar words (see O'Rourke & Holcomb, 2002).

Of particular interest in the present study is the extent to which phonological factors influence the N400. Both the latency and amplitude of the N400 have been shown to be sensitive to word-initial phonological overlap (e.g., O'Rourke & Holcomb, 2002; Connolly & Phillips, 1994; Praamstra et al., 1994). For instance, a delayed N400 is observed during sentence listening when the initial phonemes of a semantically incongruous word match the expected word (as in the *luggage*/*luck* example above; Connolly & Phillips, 1994). In addition, alliteration priming has been associated with reduced N400 for primed compared to unprimed targets (e.g., in Dutch, *beeld*–*beest*; Praamstra et al., 1994). Several studies have also illustrated that word-final phonological overlap results in a decrease in the N400 component (Coch, Grossi, Skendzel, & Neville, 2005; Boëlte & Coenen, 2002; Dumay et al., 2001; Radeau et al., 1998). In a comparison of semantic and rhyme priming, Radeau et al. (1998) demonstrated that both types of relationships led to a decreased negativity in this component. Furthermore, phonological overlap has a graded effect on the N400, with the greatest reduction for priming a whole syllable (e.g., in French, lu*rage*–ti*rage*), then rime overlap (e.g., lub*age*–tir*age*), compared to coda overlap (lubo*ge*–tira*ge*) and control primes (e.g., lusole–tirage; Dumay et al., 2001). Importantly, despite this being a phonological manipulation, this effect is arguably postlexical because it tends to be limited to real words.

Although there is some debate about whether the PMN and N400 are indeed dissociable components, or represent two parts of the same whole (Van Petten et al., 1999), taken together the pattern of findings discussed above demonstrates that the two components, at the very least, respond to dissociable types of expectancy violations in auditory word recognition. Thus, they seem appropriate for the purpose of the present study, which was to investigate the mechanisms involved in different types of phonological miscues.

The Present Study

The purpose of the present study was to investigate the role that phonological similarity plays in the time course of spoken word recognition. We used ERPs to investigate the neural underpinnings of phonological similarity effects on spoken word recognition. Of special interest were the aspects of cognitive processing that underlie cohort and rhyme effects, specifically the role of pre- and postlexical mechanisms. We used a visual-picture/spoken-word matching paradigm that was designed to reveal interactions between top–down and bottom–up processes during the time course of auditory word recognition (see Connolly, Byrne, & Dywan, 1995). This paradigm might also shed some light on the question of feedback connections from lexical to sublexical mechanisms in word recognition.

We examined how specific ERP components were differentially modulated by phonological miscues. Trials included match trials, where the spoken word matched the picture (e.g., CONE–*cone*), and three types of mismatch trials: unrelated (e.g., CONE–*fox*), rhyme (e.g., CONE–*bone*), or cohort (e.g., CONE–*comb*). It was hypothesized that these mismatch types would differentially elicit PMN and N400 in a way that can reveal distinct aspects of processing involved in disambiguating phonological similarity over the time course of spoken word recognition.

In the unrelated mismatch condition, the auditory word violates both semantic and phonological expectations. It is hypothesized that in this condition, both the PMN and N400 would be elicited. In the rhyme mismatch condition, because rhymes differ in initial phonological information from what is expected, the PMN was also expected. Although we anticipated that this condition would yield an increased N400 compared to match trials, of interest was whether the amplitude of this N400 component would be weaker than for phonologically dissimilar mismatches, reflecting the influence of the phonological expectation at the level of lexical identification. Cohort mismatches were also expected to elicit the N400 component, but not the PMN component, because the miscue does not occur until the final phoneme. For this same reason, the time course of the N400 was expected to be delayed relative to what is observed for the unrelated mismatch condition.

METHODS

Participants

A total of 15 students from the University of Western Ontario, in London, Ontario, participated in the current study (13 women, 2 men; mean age = 24 years). Each received CAN\$20 or a partial course credit for participating. All were right-handed, native English speakers with no history of hearing loss or neurological impairment. All methods and procedures were approved by the University of Western Ontario Non-Medical Research Ethics Board.

Stimuli and Procedures

Auditory stimuli were monosyllabic words spoken by an adult female English speaker, digitally recorded at 16 bits with a sampling rate of 48,828 Hz. To be compatible with our experimental presentation software (E-Prime, Psychology Software Tools, Pittsburgh, PA), the sound files were resampled to 44,100 Hz using SoundForge (Sonic Foundry, Madison, WI). Auditory stimuli were presented to the right ear using ER-3A insert earphones (Etymotic Research, Elk Grove Village, IL). Visual stimuli were color stock photographs of each object, presented on a white background using a 19-in. CRT monitor.

On each trial, a fixation cross appeared for 250 msec, following which a picture was presented. After 1500 msec, a spoken word was played while the picture remained on-screen. Participants were asked to indicate whether the picture and word matched by pressing one of two keys on a handheld keypad, pressing with their right index finger for "yes" and right middle finger for "no." There was a 1000-msec delay between the response and the beginning of next trial. Response latencies greater than 2500 msec were coded as errors (2% of trials). Participants performed six practice trials prior to the experimental task in order to familiarize them with the procedure. The experimental task consisted of 186 trials; 93 were match trials (e.g., picture: CONE; sound: *cone*), randomly interleaved with three mismatch trial conditions: unrelated mismatch (CONE: *fox*; 31 trials), cohort mismatch (CONE: *comb*; 31 trials), and rhyme mismatch (CONE: *bone*; 31 trials). Each auditory word stimulus was presented once as a match and once as a mismatch (refer to the Appendix for a list of trials, as well as frequency and neighborhood size estimates for each item). Stimulus triads (target–cohort–rhyme) were balanced for frequency as much as possible (Zeno, Ivens, Millard, & Duwuri, 1995). Each participant was randomly assigned to one of two pseudorandom stimulus sequences that counterbalanced the match versus mismatch order for each auditory word stimulus. Across the two lists, picture–word pairs were balanced so that each item appeared once as a picture and once as a word in each critical mismatch trial condition.

At the start of the experiment, participants were asked to name each of the pictures to ensure that they knew the appropriate word for each. In cases where a picture could be referred to by more than one name (e.g., saying *flower* instead of *rose*), feedback was provided to indicate the word they would hear.

Electrophysiological Recording

EEG was recorded at 500 Hz using a 64-channel cap (Quik-Caps; Neuroscan Labs, El Paso, TX) embedded with Ag/AgCl sintered electrodes, referenced to the nose tip. Impedances were kept below 5 kΩ. Electrodes were also used to record horizontal (electrodes on the outer canthi) and vertical (electrodes above and below the left eye) eye movements. Electrophysiological data were filtered on-line with a 60-Hz notch filter and off-line using a zero phase shift digital filter (24 dB, band-pass frequency: 0.1–20 Hz). Each trial was baseline corrected to the average voltage of the 100-msec prestimulus interval. Trials containing eye blinks and other artifacts were removed (determined by a maximum voltage criterion of ±75 on all scalp electrodes). Analyses were performed on the remaining trials (average nonrejected trials: 28/31 cohort, 28/31 rhyme, 29/31 unrelated, 85/93 match). ERPs were calculated from −100 to 800 msec, time-locked to the onset of the auditory word.

ERP Analyses

Analyses focused on three negative-going components commonly associated with auditory word recognition: the N100, PMN, and N400. The amplitude of each was quantified by averaging voltage values across subjects within four distinct time intervals as follows: N100, 90–110 msec; PMN, 230–310 msec; N400, 310–410 msec; and late N400, 410–600 msec (time intervals were determined based on visual inspection of the waveforms). The N100 was examined based on evidence that physical properties of auditory stimuli can influence earlier-going components, and that such effects can carry over to later components like N400 (Bonte & Blomert, 2004). The two N400 time intervals were included based on a visual inspection of the data suggesting differences across conditions that emerged at earlier versus later time periods in the N400 complex.

Statistical analyses were performed using 15 scalp sites (Fz, F3, F4, F7, F8, Cz, C3, C4, T7, T8, Pz, P3, P4, P7, P8), which provided appropriate scalp coverage to identify and differentiate the components of interest (e.g., Newman et al., 2003; Connolly & Phillips, 1994). A repeated measures analysis of variance (ANOVA) using conservative degrees of freedom (Greenhouse & Geisser, 1959) was performed on the mean amplitude at each time interval. Each ANOVA had two factors: site (15 electrodes listed above) and condition (match, cohort-mismatch, rhyme-mismatch, unrelated-mismatch). In the case of significant Site × Condition interactions, pairwise post hoc t tests were conducted to identify electrodes for which the condition-wise difference was significant.

RESULTS

Behavioral Results

Reaction time and accuracy data are listed in Table 1. Two one-way ANOVAs revealed no significant differences between the four conditions for either reaction time [$F(3, 56) = 1.06, ns$] or accuracy [$F(3, 56) = 0.64, ns$]. The data suggest that all conditions were relatively well-balanced with respect to difficulty.

Electrophysiological Results

ERP results are illustrated in Figures 1 and 2, in which each mismatch condition is separately contrasted with

Table 1. Mean (Standard Error) for Accuracy and Reaction Time (Relative to Word Onset) for Each Condition

	Accuracy (%)	*RT (msec)*
Match	95.00 (1.0)	930 (38.4)
Cohort mismatch	94.93 (1.0)	1033 (54.4)
Rhyme mismatch	95.67 (1.0)	947 (44.3)
Unrelated mismatch	96.93 (1.0)	953 (40.6)

the match condition. Analyses of N100 amplitudes (90 to 110 msec interval) revealed no interaction between site and condition [$F(42, 588) = 1.17$, ns], and no main effect of condition [$F(3, 42) = 0.32$, ns]. There was a main effect of site [$F(14, 196) = 42.82$, $p < .001$], however, this effect was anticipated because the N100 is typically strongest over central sites.

For the 230–310 msec interval, the ANOVA revealed a significant Site × Condition interaction [$F(42, 588) = 3.47$, $p < .005$], suggesting that a PMN component was being elicited by some of the conditions. This was confirmed by post hoc analyses (Table 2), which revealed increased negativity for both the unrelated mismatch and rhyme mismatch conditions over Cz and Pz compared to the match condition (Figure 1A, B and Figure 2). The PMN was not observed for cohort mismatches, consis-

tent with the initial phonological overlap between the cohort stimulus and the expected word.

Similarly, the ANOVA for the 310–410 msec time interval, corresponding to an N400 component, revealed a significant Site × Condition interaction [$F(42, 588) = 2.25$, $p < .05$]. Post hoc analyses showed that significantly greater negativity was elicited for both unrelated and rhyme mismatches over Cz and Pz (Figure 1B, C and Figure 2). An increased N400 was also observable in the cohort condition over parietal sites.

Finally, the ANOVA for the late N400 (410 to 600 msec time interval) also revealed a significant Site × Condition interaction [$F(42, 588) = 5.38$, $p < .001$]. Post hoc analyses indicated this was due to the cohort mismatch condition, which showed increased negativity in this component at frontal, central, and parietal sites (Figure 1C).
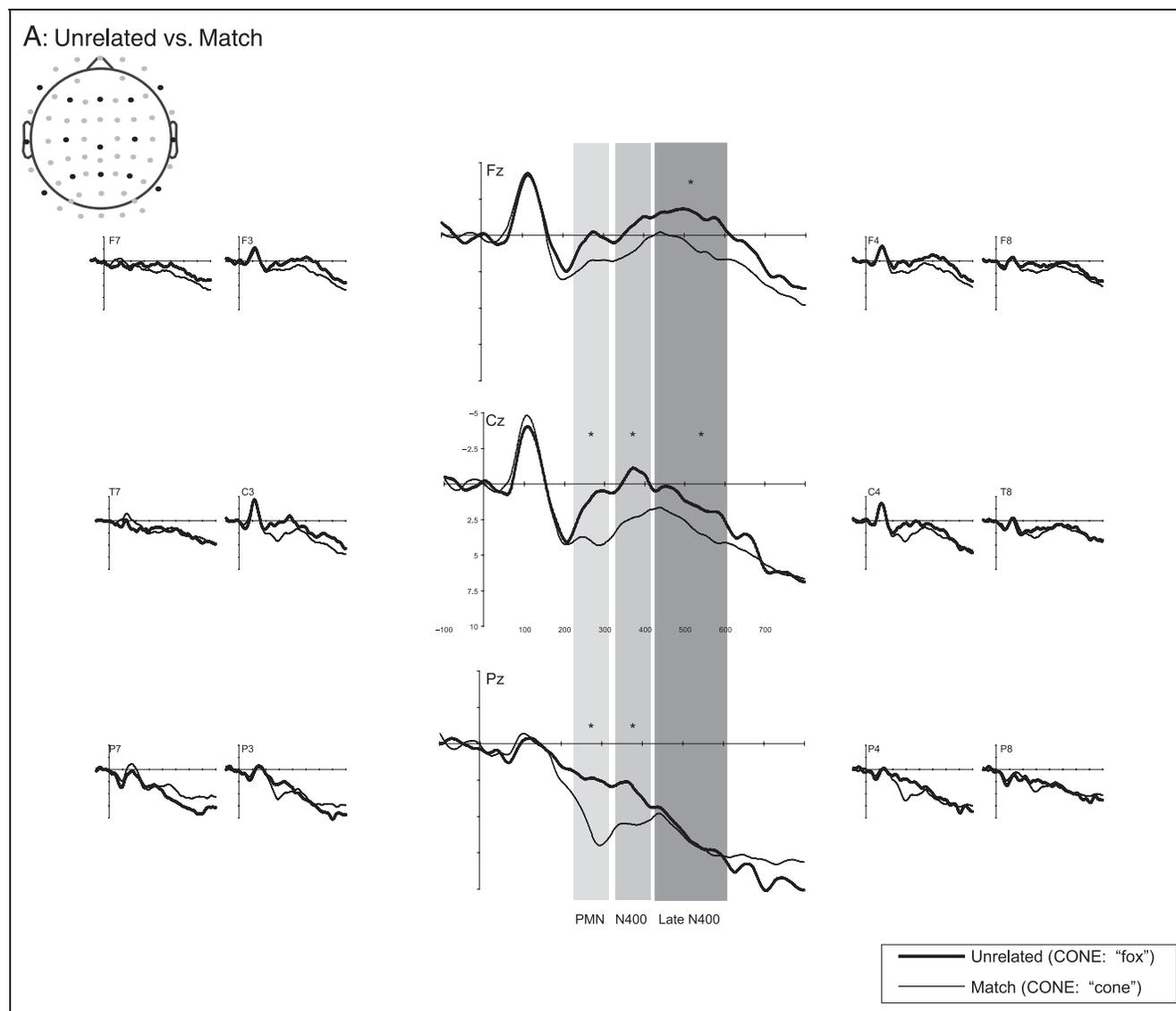
Figure 1. Average waveforms for mismatch conditions compared to the match condition. (A) Unrelated vs. Match: Results indicate a phonological mismatch negativity (PMN) and N400 effects. (B) Rhyme vs. Match: Results indicate a PMN and an N400 effect. (C) Cohort vs. Match: Results indicate a late N400 effect.
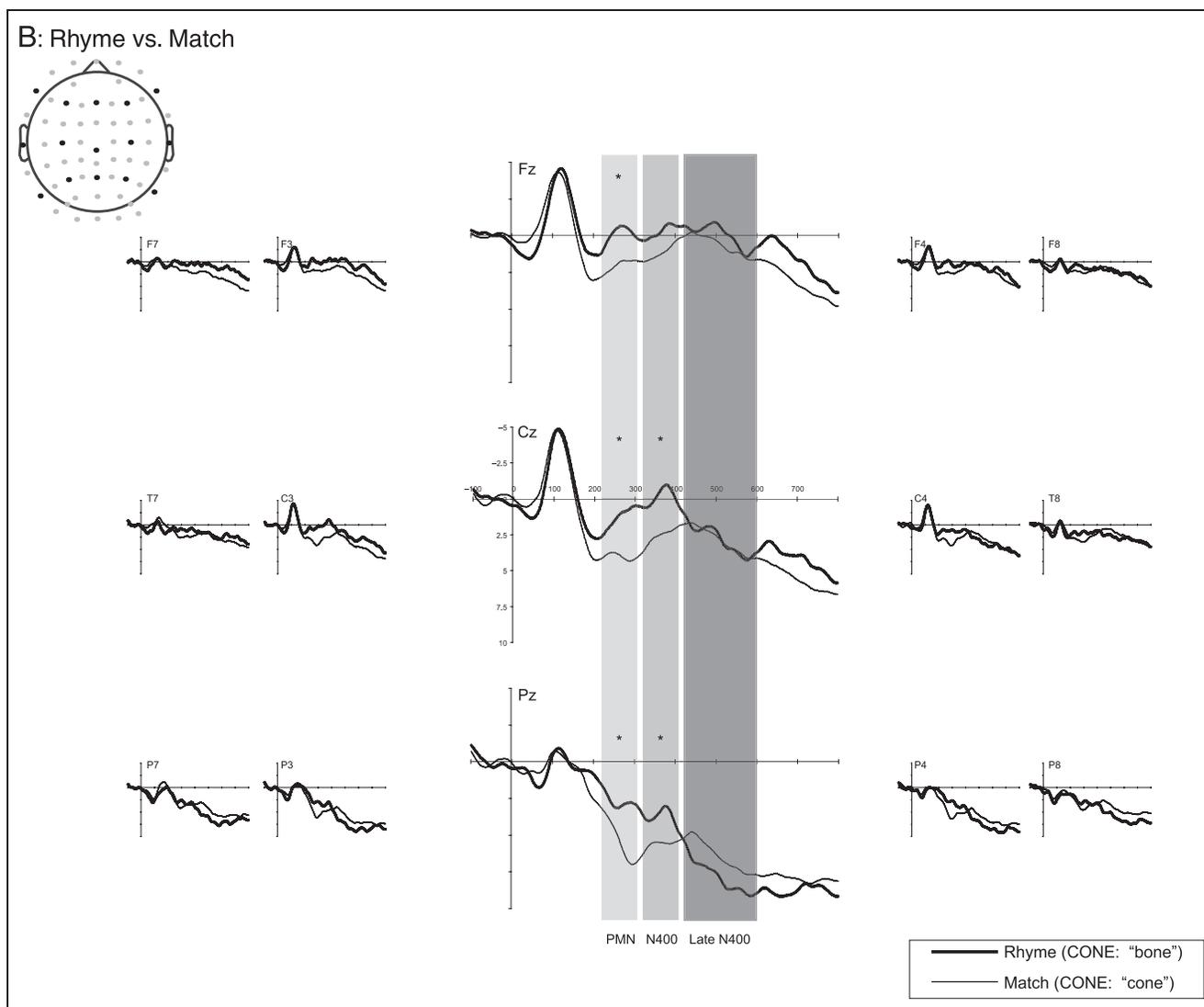
Figure 1. (*continued*)

The unrelated mismatch condition also yielded a stronger late N400 component compared to the match condition. In contrast, no such effect was observed for the rhyme mismatch condition, which did not differ significantly from the match condition at this time interval. This finding was further reinforced by an additional post hoc analysis that showed a significantly weaker late N400 effect (at Pz) for the rhyme condition compared to the unrelated mismatch condition [$t(14) = 2.01, p < .05$, one-tailed], and a stronger N400 (at Pz) for the cohort condition compared to the unrelated mismatch condition [$t(14) = -2.68, p < .01$, one-tailed]. These differences can be inferred by comparing across the scalp maps displayed in Figure 2, depicting the subtraction of each mismatch from the match condition at the three critical time intervals.

DISCUSSION

In the present study, we capitalized on the temporal sensitivity of ERPs to investigate the time course of spoken word recognition and the electrophysiological correlates of phonological competition effects. We manipulated the congruity of visually presented pictures and subsequently presented auditory words using three mismatch types: unrelated, rhyme, and cohort. Each of these violations differentially modulated the PMN and N400 components, indexing distinct influences of prelexical and lexical processes in the identification of spoken words. Consistent with prior studies, the data provide support for the suggestion that the PMN and N400 are dissociable components (Connolly & Phillips, 1994). Furthermore, as discussed below, the effects provide useful insights into understanding the basic cognitive mechanisms engaged during spoken word recognition.

The PMN is a component that is observed when the initial phonemes of a word diverge from a phonological expectation; consequently, it is proposed to reflect early influences of top–down phonological expectations on subsequent bottom–up processing of auditory inputs (Newman et al., 2003; D'Arcy et al., 2000; Connolly &

Phillips, 1994). What is different about the present study, however, is that expectations were established by displaying a visual picture prior to presenting an auditory word. Thus, the elicitation of a PMN is due to phonological expectations generated absent of a prior auditory input—instead, the phonological expectation is developed top–down, as a consequence of the visual stimulus (presumably via connections between the lexical level and the phoneme level). Consistent with our predictions, we found a negative-going component in the rhyme and unrelated mismatch conditions showing a temporal and scalp distribution consistent with a PMN. We hypothesize that this occurred for both of these conditions because the initial phoneme(s) of the auditory word mismatched the expectation. Moreover, a PMN was not observed for the cohort condition for the same reason—the onset of a cohort competitor matched the anticipated word, and as a result, no negativity was observed at this point.

Indeed, the waveforms of the cohort and match condition only began to diverge later in time.

In order to study the interaction between phoneme-level and word-specific information, we also examined how phonological similarities between an expected and perceived auditory word influenced the N400 component. This was based on the suggestion that the amplitude and latency of the N400 can reflect various aspects of lexical processing (e.g., O'Rourke & Holcomb, 2002; Hagoort & Brown, 2000; Connolly & Phillips, 1994; Praamstra et al., 1994; Holcomb & Neville, 1991; Van Petten & Kutas, 1990; Kutas & Hillyard, 1984). We observed a modulation of the N400 for both unrelated and rhyme mismatches at the early time interval, reflecting the difference between the expectation and the auditory miscue.

Of particular interest was sensitivity of the N400 component to rhyming (Coch et al., 2005; Boëlte & Coenen, 2002; Dumay et al., 2001; Radeau et al., 1998; Praamstra
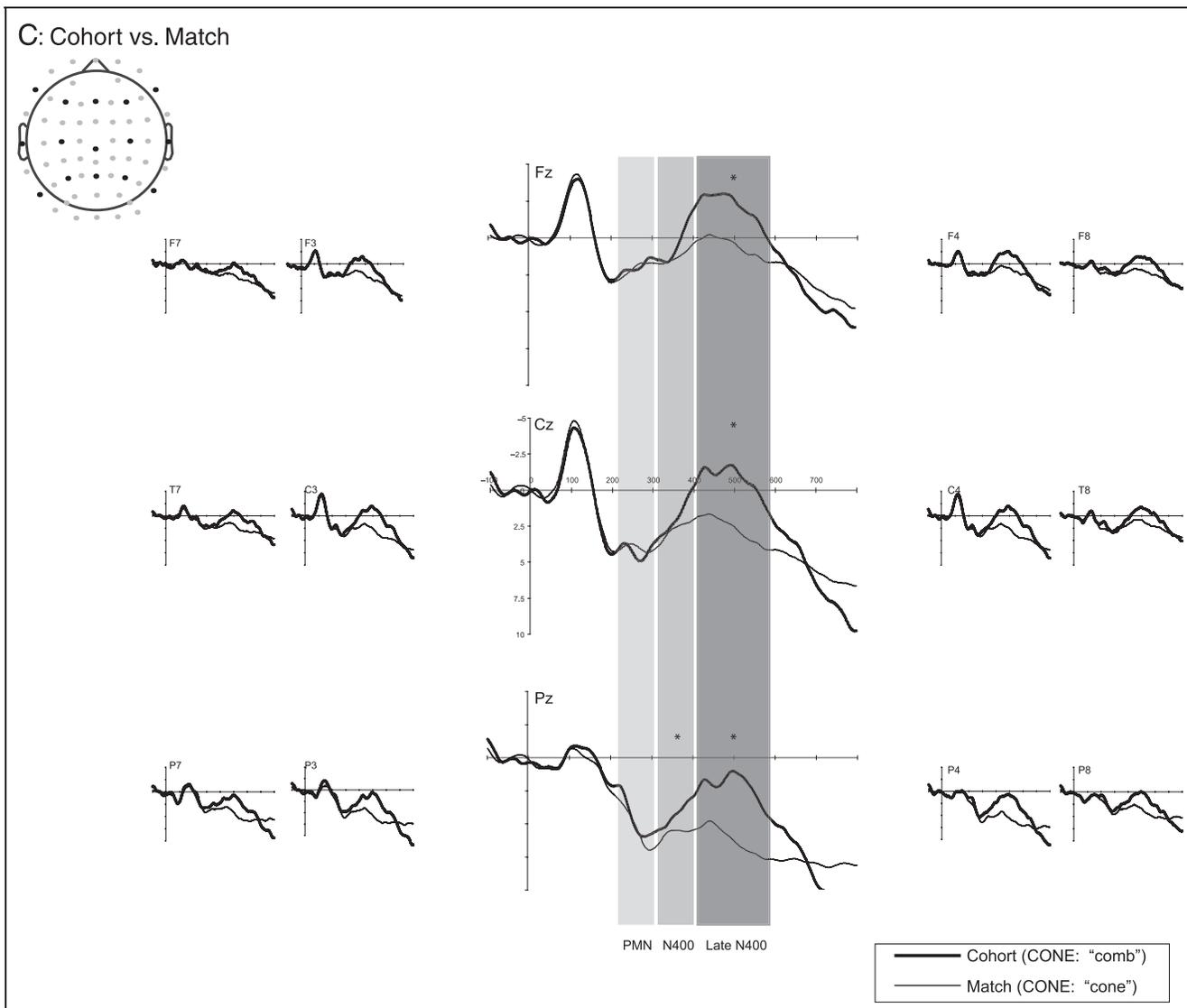
Figure 1. (*continued*)
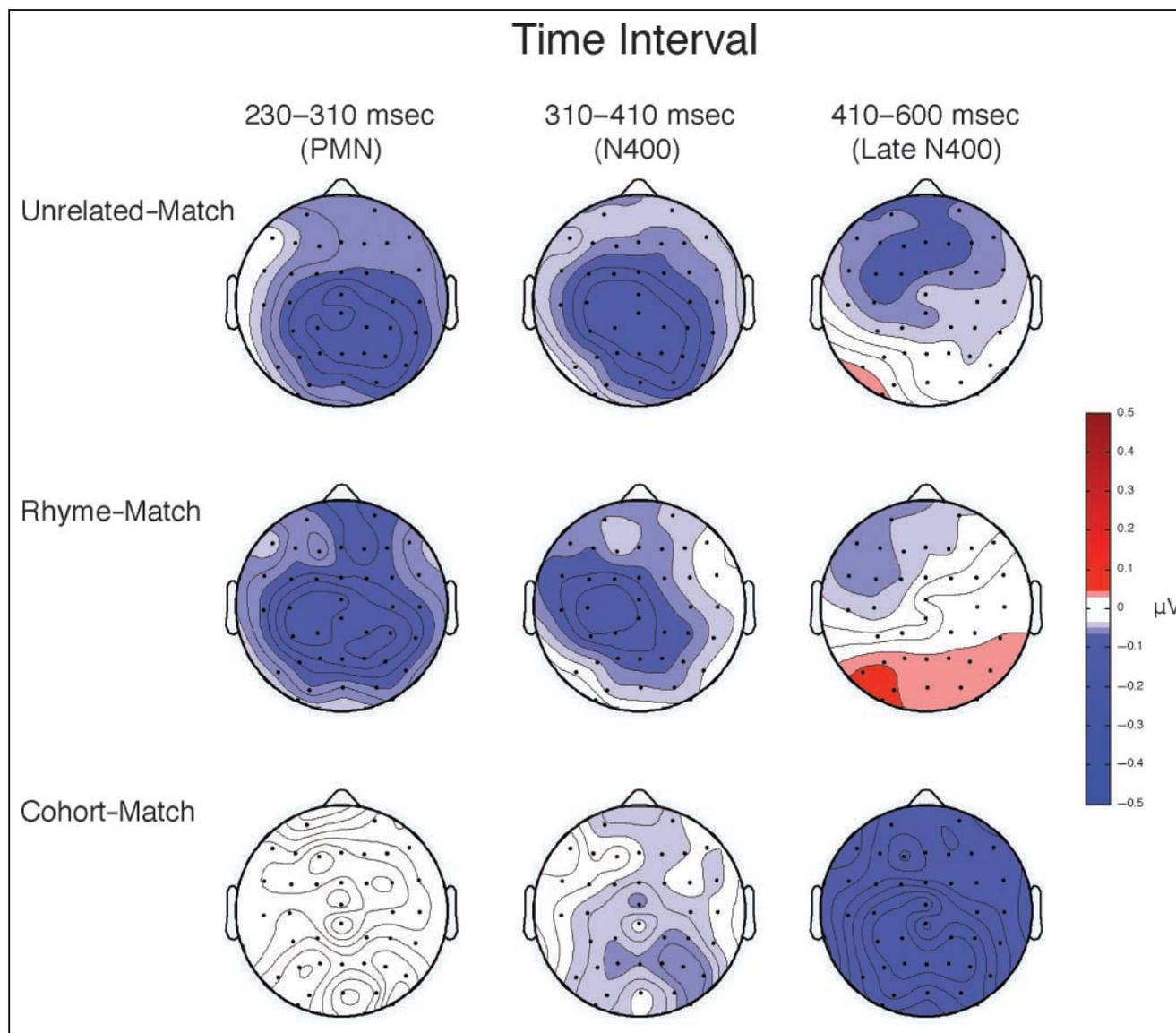
Time Interval

Figure 2. Subtraction maps illustrating the difference between the match condition and the unrelated, rhyme, and cohort conditions, respectively, computed for the three time intervals of interest. Despite early similarities in the negativity for the unrelated and rhyme mismatch conditions, there is continued negativity only for the unrelated condition in the late N400 interval. In addition, the match and cohort conditions are similar at earlier time points, but begin to diverge at the N400 period in response to the cohort mismatch.

et al., 1994). Interestingly, the N400 was more sustained for the unrelated mismatch condition compared to the rhyme mismatch condition. This was captured in our analyses by dividing the N400 wave into earlier and later time intervals. Consequently, we found that the two conditions showed differing effects in the later N400 time interval, marked by significantly lower negativity for rhyme compared to unrelated mismatches. This finding is consistent with the view that N400 amplitudes can be modulated by phonological overlap, with word-final phonological similarity (like rhymes) causing a reduction in this component (e.g., Dumay et al., 2001).

Word-initial phonological similarity also modulated the N400, providing further insight into the processes engaged in recognizing spoken words. In the cohort con-

dition, the uniqueness point of a word from a given expectation occurred later in the time course of the spoken word (e.g., cone–comb). As a result, the N400 response was shifted in time such that the increased negativity was only observed at the later time interval. We attribute the later timing to the fact that the miscue occurred later in the word, rather than at the onset. This finding suggests that initial phonological expectations do, in fact, influence spoken word identification, and furthermore, that lexical interpretations are being continuously updated as more phonological information becomes available. Similar effects of word-initial overlap on the timing of the N400 have been observed in prior studies (O'Rourke & Holcomb, 2002; Connolly & Phillips, 1994). This late N400 was proposed to reflect the delay in

Table 2. Comparison of Match vs. Mismatch Conditions for PMN, N400, and Late N400

Component	Electrode	t
PMN		
Cohort vs. Match	Fz	0.11
	Cz	−0.52
	Pz	0.81
Rhyme vs. Match	Fz	1.99*
	Cz	3.19**
	Pz	4.03***
Unrelated vs. Match	Fz	1.51
	Cz	3.10**
	Pz	3.82***
N400		
Cohort vs. Match	Fz	1.44
	Cz	1.09
	Pz	1.98*
Rhyme vs. Match	Fz	1.59
	Cz	2.70**
	Pz	1.91*
Unrelated vs. Match	Fz	1.58
	Cz	2.68**
	Pz	3.07***
Late N400		
Cohort vs. Match	Fz	3.90***
	Cz	1.81*
	Pz	4.17***
Rhyme vs. Match	Fz	0.79
	Cz	0.81
	Pz	−1.73
Unrelated vs. Match	Fz	1.95*
	Cz	1.77*
	Pz	0.30

*$p < .05$ (one-tailed).

**$p < .01$ (one-tailed).

***$p < .001$ (one-tailed).

the miscue and the temporal nature of the lexical identification process; however, as we will discuss subsequently, it may also reflect somewhat different aspects of processing (see Kujala, Aljo, Service, Ilmoniemi, & Connolly 2004; Connolly, Service, D'Arcy, Kujala, & Alho, 2001).

Cohort mismatches modulated the N400 in a way that appeared to be larger in magnitude than any of the other mismatch effects. This effect was marked by a large negativity across the late N400 time interval, which was significantly larger than was observed for unrelated mismatches. We consider two possible interpretations for this result. The first is that, in the cohort condition, the initial phonological input incorrectly confirms the expectation generated by the picture, strengthening that expectation; however, at the final phoneme, a mismatch occurs and the interpretation must be restructured. These early misleading effects, coupled with competition effects, may require a more effortful resolution of the mismatch. Thus, compared to unrelated mismatches where the bottom–up information is not misleading, a larger negativity is observed for cohorts.

A second interpretation is that the larger apparent N400 for cohorts is, in fact, an additive effect of a late PMN and a late N400. In this condition, the simultaneous mismatch in phonemic and semantic information occurs later in the time course of the auditory word. It could be that the large negativity reflects some additivity in the two processes; however, this alternate interpretation may only hold in the case of expectancy generation or priming paradigms and may not generalize to our understanding of spoken word recognition as well as the former interpretation. Under this assumption, the unique contribution of each of these processes was difficult to disentangle in this condition due to the timing of the mismatch: Both stimulus duration and the uniqueness point varied across stimulus items, factors that are difficult to control for in natural speech. Thus, the underlying components reflecting these effects may have merged in such a way that they were difficult to isolate (Kujala et al., 2004; Connolly et al., 2001). Although others have pointed out that the uniqueness point can impact the latency of the N400, on the present task this was not controlled for, making it difficult speak to this particular question here. Future investigations could address this by more systematically manipulating and measuring the uniqueness point of auditory words. These interpretations are not mutually exclusive, nor do they undermine the more general point of our analyses; specifically, the observed N400 modulation observed for both cohort and rhyme mismatches reveals that phonologically driven lexical competition occurs in an on-line fashion during spoken word recognition, and the timing of these mismatches influences the pattern of results in important ways (Magnuson et al., 2007).

Evidence for Continuous Mapping Models of Spoken Word Recognition

The present study provides neurophysiological support for models suggesting that spoken language is processed in a continuous and dynamic fashion (e.g., TRACE). The

observed phonological competition effects provide useful insights into the mechanisms responsible for processing speech. In TRACE, acoustic information activates corresponding phonological information, which in turn activates word-specific ("lexical") information. As phonemic information unfolds, these mechanisms operate to identify the input; at the same time, expectations derived from lexical cues such as frequency, and top–down cues such as context, help to constrain or guide this process. TRACE proposes that both prelexical and lexical mechanisms are engaged in a nonlinear and interactive fashion as listeners recognize spoken words. In this model, competition occurs via lateral inhibition of lexical-level units that correspond to individual word identities. Because of interaction between the phoneme level and the lexical level, this inhibition is strengthened for phonologically similar items, causing phonological competition effects. The assumption of feedback is a key prediction of this model, and distinguishes it from models that assume a more linear bottom–up process (e.g., Cohort and also Shortlist), or which do not encode temporal information at all (e.g., NAM, Merge).

Consider the following account of this paradigm: On any given trial, when a picture is presented (e.g., CONE), the representation of that concept is activated at the lexical level, with lateral inhibition acting to suppress activation of all other lexical-level units. However, top–down connections between the lexical-level and phoneme-level result in the activation of the phonological units corresponding to this word (/k/, /o/, and /n/). Subsequently, activation that feeds from the phonological level to the lexical level not only reinforces the expectation of CONE but also results in the partial activation of lexical-level units that are phonologically similar (e.g., cohorts and rhymes like COMB and BONE). Next, the auditory cue is presented, which corresponds to bottom–up acoustic information that activates the corresponding phoneme-level, and in turn, lexical-level units.

On the *match* condition, this bottom–up information confirms the prior expectations, and thus, competitors at the lexical level are easily eliminated. However, for mismatches, expectancy violation effects are revealed due to bottom–up inputs that are inconsistent with activation incurred at the phoneme and/or word levels. The differential recruitment of the PMN and N400 components informs us about how this competition occurs. In contrast, for *unrelated mismatches* (CONE–*fox*), the auditory word violates the expectation completely, both at the phoneme and word levels. Thus, both components of interest are elicited, reflecting competition between the expected and perceived word at the phonological and lexical levels, respectively.

For *rhyme mismatches* (CONE–*bone*), the auditory word violates the expectation at word onset. The prior visual input CONE created the expectation of /k/, which mismatches the /b/ onset, yielding a PMN. Although this initial violation is quickly perceived, lexical-level in-

fluences of rhyme similarity are observed. Seeing CONE activates the /o/ and /n/ phoneme units, which have reciprocal connections to word units like *bone*. On the rhyme condition, because the rhyme competitor (e.g., *bone*) is already activated, its recognition is facilitated—as indicated by the reduction in the later N400 component.

For *cohort mismatches*, the onset of the auditory word overlaps with the phonological expectation (e.g., /kom/) and initially confirms that expectation. The bottom–up inputs match the anticipated /k/ and /o/ phoneme units, further strengthening the activation of CONE which further serves to inhibit competitors (including *comb*). Correctly recognizing the word when the mismatch finally occurs requires relatively more effort because it involves both activating COMB and deactivating the expected CONE. This process of rejecting the expectation and accepting the competitor is reflected by the larger and later occurring N400 component than if the bottom–up mismatch had occurred earlier.

As this account indicates, violations are being processed at both the phoneme level and the lexical level, indexed by the PMN and N400, respectively. The suggestion is these two components differ in more than just latency, and indeed, reflect subtly different processes in word recognition. In addition, although a strictly feedforward theory would suggest that prelexical processes are those engaging phonemic processing (as indicated by the PMN), and lexical processes are those engaging word-specific knowledge (as indicated by the N400), our data suggest that these are engaged in an interactive fashion. We did observe the influence of lexical expectations on on-line processing of prelexical phonological inputs, as indicated by the elicitation of the PMN for unrelated and rhyme mismatches, and also by the later timing of the N400 for cohort mismatches. However, we also found an influence of phonology at the lexical level, revealed by the modulation of the late N400 component, which was reduced for rhymes and increased for cohorts. Taken together, the later occurrence and the increased N400 amplitudes for cohorts supports the position that there are temporally mediated influences of onset similarity that not only result in the activation of a competitor set but also drive lateral inhibition among initially similar words. This is an explanation that may account for the relative strength of the cohort competition effect observed in previous behavioral investigations (e.g., Desroches, Joanisse, & Robertson, 2006; Allopenna et al., 1998; Marslen-Wilson & Zwitserlood, 1989). The observed N400 rhyme effects strongly suggest that there is an interaction between the phoneme and lexical levels of representation. Given that the reduction in the N400 component suggests the prior activation of the auditory word, the present effects can only be accounted for by a model that allows for both top–down and also bottom–up connections between levels of representation. That is, it is difficult to conceptualize how this effect could occur in a

model that does not include re-entrant connections from a lexical/semantic to a phonological processing layer. This is a finding relevant to the ongoing debate of feedback in spoken word recognition, or at the very minimum, the on-line influence of top–down information during recognition (e.g., Magnuson et al., 2005; Norris et al., 2000).

Conclusions

The present findings provide evidence for the underlying mechanisms engaged during spoken language processing, supporting dynamic models of spoken word recognition such as TRACE. Although other models (i.e., Shortlist/Merge) can account for bottom–up phonological competition in a satisfactory way, the present findings illustrate interaction between levels, such that top–down word-level expectations influence how phoneme-level information is processed. Thus, a novel aspect of this study is specifically that expectancies are generated using pictures rather than auditory words (note that Lupker & Williams, 1989 have also observed behavioral rhyme priming effects using a similar paradigm). Pictures are used to activate lexical-level representations, which in turn activate phonological representations in a top–down fashion. This is quite different from auditory priming, where similar effects could be explained strictly via residual phoneme-level activation (see Praamstra et al., 1994). The observation of a PMN for both unrelated and rhyme mismatches supports the claim that such top–down (word–phoneme level) connections are being used as speech unfolds, and in a way that can only be accounted for by a model that emphasizes the temporal structure of processing. Moreover, the observed influence of rhyme similarity on the N400 can only be accounted for by a model that assumes interaction between the phoneme and lexical levels of representation, which we believe can only be explained by top–down feedback connections between levels. Models that are strictly feedforward can only account for lexical influences at a postlexical decision stage, or via residual activation of a previously presented auditory prime. Neither of these can account for the current data, which illustrate the on-line influence of lexical expectancies on phonological processing derived from a previously presented visual—rather than auditory—cue. Thus, the picture–word paradigm allowed us to address broad questions of auditory word processing, offering findings that help to disentangle the influences of prelexical and lexical information in the recognition of spoken words.

APPENDIX

Stimulus Triads

Frequencies (Zeno et al., 1995; on a scale of parts per million) and number of neighbors (Davis, 2005) are indicated in parentheses.

Picture		Cohort		Rhyme		Unrelated
rose	(1174, 41)	road	(2356, 34)	hose	(95, 33)	sock
cake	(520, 24)	cage	(348, 14)	rake	(47, 260)	hose
corn	(895, 29)	cord	(221, 36)	horn	(259, 25)	mat
bat	(224, 34)	bath	(217, 15)	mat	(115, 32)	wheel
cart	(203, 25)	card	(549, 33)	dart	(40, 15)	soap
mouse	(1864, 12)	mouth	(941, 3)	house	(9448, 21)	rake
peach	(73, 18)	peas	(125, 21)	beach	(764, 21)	ghost
stone	(1662, 12)	stove	(372, 9)	phone	(473, 26)	dart
clock	(624, 17)	cloth	(828, 4)	block	(731, 12)	bone
toast	(106, 12)	toes	(247, 34)	ghost	(309, 12)	house
lock	(243, 31)	log	(397, 24)	sock	(38, 20)	beach
boat	(1589, 26)	bowl	(420, 27)	coat	(791, 28)	knife
cone	(118, 28)	comb	(118, 22)	bone	(538, 27)	block
hat	(1001, 36)	hand	(5538, 14)	cat	(1772, 32)	bun
rope	(749, 26)	robe	(83, 20)	soap	(228, 21)	jet

APPENDIX (*continued*)

Picture		Cohort		Rhyme		Unrelated
seal	(274, 23)	seed	(350, 32)	wheel	(776, 24)	tape
knot	(69, 25)	knob	(49, 18)	pot	(536, 28)	cane
bug	(173, 27)	bun	(9, 30)	mug	(18, 23)	kite
note	(1017, 22)	nose	(1187, 28)	goat	(226, 19)	ship
cape	(233, 21)	cane	(201, 31)	tape	(430, 14)	goat
knight	(103, 22)	knife	(415, 9)	kite	(240, 22)	horn
chip	(143, 19)	chick	(141, 18)	ship	(1759, 17)	cat
suit	(526, 22)	soup	(328, 20)	boot	(67, 26)	fan
net	(359, 24)	neck	(944, 14)	jet	(204, 20)	pot
purse	(99, 20)	pearl	(117, 23)	nurse	(362, 13)	bell
doll	(174, 17)	dog	(3571, 16)	ball	(2433, 28)	match
wing	(248, 21)	wig	(25, 23)	king	(2747, 16)	ball
map	(1377, 22)	cap	(474, 29)	match	(506, 18)	fox
shell	(477, 18)	shed	(234, 21)	bell	(865, 24)	mug
can	(34823, 26)	cab	(78, 16)	fan	(127, 21)	nurse
box	(2030, 21)	bomb	(107, 16)	fox	(654, 17)	king

Note

1. The PMN should not be confused with the similarity named mismatch negativity, a distinct ERP component not examined in the present study.

REFERENCES

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419–439.

Boëlte, J., & Coenen, E. (2002). Is phonological information mapped onto semantic information in a one-to-one manner? *Brain and Language, 81,* 384–397.

Coch, D., Grossi, G., Skendzel, W., & Neville, H. (2005). ERP nonword rhyming effects in children and adults. *Journal of Cognitive Neuroscience, 17,* 168–182.

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language, 32,* 193–210.

Connolly, J. F. (in press). Event related potentials and magnetic fields associated with components and subcomponents that enable spoken word recognition. In M. Spivey, M. F. Joanisse, & K. McRae (Eds.), *Cambridge handbook of psycholinguistics*. Cambridge: Cambridge University Press.

Connolly, J. F., Byrne, J. M., & Dywan, C. A. (1995). Assessing adult receptive vocabulary with event-related potentials: An investigation of cross-modal and cross-form priming. *Journal of Clinical and Experimental Neuropsychology, 7,* 548–565.

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

Connolly, J. F., Phillips, N. A., Stewart, S. H., & Brake, W. G. (1992). Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. *Brain and Language, 43,* 1–18.

Connolly, J. F., Service, E., D'Arcy, R. C. N., Kujala, A., & Alho, K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *Cognitive Neuroscience and Neuropsychology, 12,* 237–243.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16,* 507–534.

D'Arcy, R. C. N., Connolly, J. F., & Crocker, S. F. (2000). Latency shifts in the N2b component track phonological deviations in spoken words. *Clinical Neurophysiology, 111,* 40–44.

Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods, 37,* 65–70.

Desroches, A. S., Joanisse, M. F., & Robertson, E. K. (2006). Specific phonological impairments in dyslexia revealed by eyetracking. *Cognition, 100,* B32–B42.

Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioural and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience, 13,* 121–143.

Frauenfelder, U. H., & Tyler, L. K. (1987). The process of spoken word recognition: An introduction. *Cognition, 25,* 1–20.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24,* 95–112.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech: Semantic ERP effects. *Neuropsychologia, 38,* 1518–1530.

Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19,* 286–300.

Kujala, A., Alho, K., Service, E., Ilmoniemi, R. J., & Connolly, J. F. (2004). Activation in the anterior left auditory cortex associated with phonological analysis of speech input: Localization of the phonological mismatch negativity response with MEG. *Cognitive Brain Research*, *21,* 106–113.

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

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19,* 1–36.

Lupker, S., & Williams, B. A. (1989). Rhyme priming of pictures and words: A lexical activation account. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1033–1046.

Magnuson, J. S., Dixon, J. A., Tanenhaus, M. D., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science, 31,* 1–24.

Magnuson, J. S., Strauss, T., & Harris, H. (2005). Interaction in spoken word recognition models: Feedback helps. *Proceedings of the XXVII Annual Meeting of the Cognitive Science Society* (pp. 1379–1384). Mahwah, NJ: Erlbaum.

Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word recognition and learning: Studies with artificial lexicons. *Journal of Experimental Psychology: General*, *132,* 202–227.

Marslen-Wilson, W., Moss, H. E., & van Halen, S. (1996). Perceptual distance and competition in lexical access. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1376–1392.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8,* 1–71.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 576–585.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1–86.

Münte, T. F., Say, T., Clahsen, H., Schiltz, K., & Kutas, M. (1999). Decomposition of morphologically complex words in English: Evidence from event-related brain potentials. *Cognitive Brain Research, 7,* 241–253.

Newman, R. L., Connolly, J. F., Service, E., & McIvor, K. (2003). Influence of phonological expectations during a phoneme deletion task: Evidence from event-related brain potential. *Psychophysiology, 40,* 640–647.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52,* 189–234.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23,* 299–370.

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

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

Radeau, M., Besson, M., Fonteneau, E., & Castro, S. L. (1998). Semantic, repetition and rime priming between spoken words: Behavioural and electrophysiological evidence. *Biological Psychology, 48,* 183–204.

van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience, 13,* 967–985.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 394–417.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition, 18,* 380–393.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40,* 374–408.

Zeno, S. M., Ivens, S. H., Millard, R. T., & Duwuri, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Sciences, Inc.