Title: Electrophysiological markers of pre-lexical speech processing: Evidence for bottom–up and top–down effects on spoken word processing

Journal: Biological Psychology

Author(s): Newman, R. L., & Connolly, J. F.

Year: 2009

Version: Post-Print

Original Citation: Newman, R. L., & Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: Evidence for bottom-up and top-down effects on spoken word processing. Biological Psychology, 80(1), 114-121. https://doi.org/10.1016/j.biopsycho.2008.04.008

Electrophysiological markers of pre-lexical speech processing: Evidence for bottom–up and top–down effects on spoken word processing

Newman, R. L.[a, *], & Connolly, J. F.[b]

[a]Department of Psychology, Acadia University, 18 University Avenue, Wolfville, NS B4P 2R6, Canada

[b]Department of Linguistics and Languages, McMaster University, Canada

* Corresponding author. Tel.: +1 902 585 1405; fax: +1 902 585 1078. E-mail address: randy.newman@acadiau.ca (R.L. Newman).

Abstract

The present study was designed to investigate the electrophysiological consequences of a mismatch between initial phoneme expectations and the actual spoken input. Participants were presented with a word/nonword prompt with the instruction to delete the initial sound (e.g., *snap without the /s/*; *snoth without the /s/*) and determine the resulting segment. Following the prompt, an aurally presented response that matched/mismatched expectations (e.g., *nap/tap*; *noth/toth*) was presented. The Phonological Mapping Negativity (PMN), a response associated with phonological processing, was largest to mismatching responses, and was not dependent on the lexical status of response items. An N400-like response was also largest to mismatching responses; however, in contrast to the PMN, the N400-like response differentiated mismatching words from mismatching nonwords. These findings highlight a functional dissociation between the PMN and N400, and establish the PMN as a neural marker representing the goodness-of-fit between initial phoneme expectations and the actual spoken input.

Keywords

Phonological Mapping Negativity (PMN); N400; Phonological processing; Event-related brain potentials; Phoneme deletion; Phonological mismatch negativity; Evoked potentials

1. *Introduction:*

Models of spoken word recognition typically incorporate the idea that multiple lexical candidates are activated in the course of speech perception, and that selection of a target candidate occurs once a match is achieved between sensory input and a lexical representation (Marslen-Wilson, 1987; Norris, 1994; McClelland and Elman, 1986). Models differ, however, with respect to how lexical candidates become activated and how mismatching candidates become eliminated from the competitor set. In the original cohort model (Marslen-Wilson, 1987), the candidate set is determined on the basis of the initial speech input (i.e., bottom–up input), and candidates are eliminated based on a mismatch in speech input. For instance, the word candidate *captain* would be eliminated from the cohort at the second vowel of the input captive. In contrast to the original cohort model, the TRACE model does not require a strict match between speech input and lexical representations (McClelland and Elman, 1986). TRACE allows for the continuous mapping of speech input onto lexical representations, such that activation of lexical representations is based on the degree of overlap between speech input and the activated set of word candidates. The consequence of this more graded approach, for example, is that the speech input *beaker* will activate overlapping representations, including those that do not share similar onsets, such as speaker. A target word is eventually selected by means of lateral inhibition between competitors. In the case of the heard input *beaker*, the lexical representation *BEAKER* will achieve a higher level of activation due to its match with the speech input, and thus, will inhibit the activation level of competitors, such as *SPEAKER*, via lateral inhibition.

In order to tease apart these competing theories of word recognition, investigations have focused on whether words that rhyme with a target representation become activated. The original cohort model would predict not, since it requires a strict match between input and the target representation. TRACE, on the other hand, would predict that rhyming words would gradually become weakly activated as input is mapped onto lexical representations. The results have been mixed. Studies employing cross-modal priming have shown evidence for the activation of rhyming words, but only when the initial onset of the rhyme competitor (e.g., *bear*) shares phonetic features with the speech input (e.g., *pear*; Connine et al., 1993; Marslen-Wilson et al., 1996). However, evidence from eye movement studies suggests that rhyming effects may be underestimated, such that even rhyme competitors (e.g., *speaker*) whose onsets are not similar in features to that of the target candidate (e.g., *beaker*) become activated as speech input unfolds (Allopenna et al., 1998).

Electrophysiological techniques, such as event-related brain potentials (ERP), offer a valuable approach for studying the microstructure of spoken word recognition. In particular, two ERP components appear sensitive to differing aspects of spoken word recognition: the N400, a component traditionally associated with semantic processing of spoken or written words (Bentin et al., 1993; Kutas and Hillyard, 1980), and an earlier occurring negativity, the Phonological Mapping Negativity (PMN),[1] which has been previously linked to phonological processes (Connolly et al., 1992, 1990). Though the N400 is typically associated with semantic analysis, there is a sizable literature showing that the N400 is also modulated by phonological factors (Dumay et al., 2001; Praamstra and Stegeman, 1993; Praamstra et al., 1994; Radeau et al., 1998; Rugg, 1984a,b). The phonological N400 which has been referred to as the N450 and as being N400-like, shares a similar latency and topography as the semantic N400, and thus the two responses are often considered as being one in the same (Kutas and Van Petten, 1988; Praamstra et al., 1994); although others (Connolly et al., 1995) have suggested that at least in the Praamstra research, the phonological effect may be more attributable to an earlier response (occurring in the 300 ms area) than to the later "N400" response. Phonological priming effects, like semantic congruity effects, have been found to be larger over parietal regions (Dumay et al., 2001; Praamstra et al., 1994; Rugg, 1984a,b). While phonological priming effects in the visual modality have shown evidence of a right hemispheric asymmetry (Rugg, 1984a,b), those in the auditory modality have exhibited equivalent hemispheric distributions (Dumay et al., 2001).

Studies employing phonological priming paradigms have shown the N400 to be smaller in amplitude for prime-target pairs that rhyme relative to non-rhyming pairs (Dumay et al., 2001; Praamstra and Stegeman, 1993; Praamstra et al., 1994; Radeau et al., 1998; Rugg, 1984a,b).

---

[1] *The Phonological Mapping Negativity (PMN) was referred to as an "N200" in the earliest studies and was subsequently labelled the Phonological Mismatch Negativity by Connolly and Phillips (1994) when its functional importance was better appreciated. This label has, however, caused some understandable confusion related to the classic Mismatch Negativity (MMN) with which it shares very few features. As a consequence, the PMN now refers to the Phonological Mapping Negativity which more accurately captures its functional relationship to phonological processing generally rather than to mismatching processes only.*

Furthermore, the amplitude of the N400-like response has been shown to decrease linearly as the degree of phonological overlap between prime-target pairs increases (Dumay et al., 2001). Dumay et al. (2001) varied the degree of phonological overlap between aurally presented prime target pairs. Related target stimuli shared the last syllable, rime, or coda with primes, while unrelated targets had no overlap with primes, and thus served as control items. The phonological priming effect (i.e., reduction in N400-like amplitude) was greatest for targets in the syllable overlap condition, intermediate for the rime overlap condition, and smallest for the coda overlap condition, which itself did not differ from the unrelated condition. Interestingly, the phonological priming effect was further modulated by the lexical status of the target. Smaller phonological priming effects (i.e., larger N400-like amplitude), particularly for targets in the rime condition, were observed for nonword targets compared to word targets. Priming effects which were dependent on target lexicality were proposed to reflect lexical selection mechanisms, whereas those effects observed for both word and nonwords targets were proposed to reflect pre-lexical mechanisms (Dumay et al., 2001).

Unlike the N400, the PMN has been proposed by Connolly and colleagues to reflect mechanisms operating entirely at the prelexical level (Connolly and Phillips, 1994; Connolly et al., 2001; Newman et al., 2003). For example, Connolly and Phillips (1994) measured ERP responses as participants were presented with auditory sentences in which the terminal word varied in semantic and/or phonological constraint. In the fully congruent condition, sentences ended with the high cloze probability word for that sentence (e.g., *The piano was out of tune*.). In the phonological condition, the terminal word shared its initial phoneme sequence with that of the highest close probably ending for that sentence, but was semantically inappropriate (e.g., *The gambler had a streak of bad luggage [luck]*). In the semantic condition, the terminal word was semantically congruent, but differed phonologically from the high cloze probability ending for the sentence (e.g., The pig wallowed in the pen [*mud*].). Finally, in the fully incongruent condition, the terminal word was semantically inappropriate and its initial phoneme sequence differed from that of the high cloze probability word (e.g., *Joan fed her baby some warm nose [food]*). In the two critical conditions, the phonological and the semantic, a clear PMN in the relative absence of an N400 was observed in the phonological condition, while an N400 in the relative absence of a PMN was apparent in semantic condition.

Connolly and Phillips (1994) concluded that the PMN and N400 represent distinct mechanisms of spoken word recognition. The former reflects a mismatch between bottom–up speech input and an activated lexical representation, whereas the later occurring N400 reflects top–down influences on word recognition. For example, in the sentence, *The pig wallowed in the pen*, the speech input pen elicits a PMN due to a mismatch in expected speech onset (e.g., *mud*); however, the lexical representation for PEN is still activated, as denoted by the relative absence of the N400, due to its semantic-fit with the sentence context.

Consistent with their proposed distinctiveness, the PMN and N400 are sensitive to different sized units of phonological information. In Newman et al. (2003), participants were instructed to delete the initial consonant from an aurally presented CCVC word (e.g., *Clap, /k/*). The prime word was followed by presentation of the correct choice (e.g., *lap*) or the incorrect choice that contained a mismatch at the initial phoneme (e.g., *cap, nose*). Unlike the N400, which was modulated by large units of phonological overlap (e.g., rime overlap), the PMN was tuned specifically to mismatches in the initial phoneme. That is, while the N400 was largest to incorrect responses that did not exhibit rime overlap with the anticipated answer (e.g., *nose*), the PMN did not distinguish between incorrect choices that shared the same rime as the expected response (e.g., *cap*) and those that were completely phonological unrelated to the anticipated response (e.g. *nose*). On the basis of their findings, Newman et al. (2003) proposed that the presence of a single mismatch in bottom–up input is sufficient to elicit the PMN; a process that was likened to the cohort model's intolerance to mismatches between input and a target representation. Furthermore, the PMN appears to represent the outcome of an autonomous matching process (Newman et al., 2003). That is, the matching process is based on comparing bottom–up input to a target representation, and is not subject to top–down influences from the potential set of competitors (e.g., rhyming words, semantically related words).

Alternative interpretations of the PMN have been offered (Hagoort and Brown, 2000; Van Petten et al., 1999; van den Brink et al., 2001). In an experiment similar to Connolly and Phillips (1994), Hagoort and Brown (2000) presented participants with spoken sentences in which the terminal ending was either semantically congruous or semantically incongruous, and whose initial phoneme did not match that of the appropriate ending. Two negativities were observed to semantically inappropriate endings; a response the authors equated with the PMN but termed the N250, and the N400 response. Hagoort and Brown (2000) argued that the N250 represents the activation of early lexical selection processes in which word candidates derived from the acoustic input are compared with semantic expectations based on the sentential context. A similar explanation for the processing nature 1 The Phonological Mapping Negativity (PMN) was referred to as an "N200" in the earliest studies and was subsequently labelled the Phonological Mismatch Negativity by Connolly and Phillips (1994) when its functional importance was better appreciated. This label has, however, caused some understandable confusion related to the classic Mismatch Negativity (MMN) with which it shares very few features. As a consequence, the PMN now refers to the Phonological Mapping Negativity which more accurately captures its functional relationship to phonological processing generally rather than to mismatching processes only. R.L. Newman, J.F. Connolly / Biological Psychology 80 (2009) 114–121 115 of the PMN was offered by van den Brink et al. (2001). A more extensive review of these differing accounts is given in Newman et al. (2003).

In order to clarify the functional nature of the PMN,Connolly et al. (2001) designed a phonological priming study assessing whether the lexical status of a target representation influenced the PMN. They hypothesized that if the PMN represented early lexical selection, then its amplitude and/or latency would be differentially influenced by the lexical status of the auditory targets. Participants were presented with a visual word/nonword (e.g., *house*/*telk*) followed by a letter (e.g., *m*/*w*), and were told to anticipate the auditory word/ nonword that would begin with the letter and would rhyme with the visual word/nonword. Thus, matching trials consisted of auditory targets that had been phonologically primed (e.g.,[2] *house, m: mouse or telk, w: welk*), whereas mismatching trials consisted of auditory targets that had not been phonologically primed (e.g., *house, m: barn*). The PMN proved to be insensitive to the lexicality of auditory targets. While the PMN was largest to mismatching trials, there was no significant difference in the amplitude of the PMN between mismatching word and nonword trials. Thus, the authors concluded that the PMN is not influenced by the semantic nature of the competitor environment, and instead represents an autonomous stage of phonological analysis in which speech input is matched to a target representation derived by task demands. However, MEG data collected at the same time found a substantial but non-significant difference in PMN latencies between words and nonwords preventing any firm conclusions about the nature of the PMN based on this paradigm (Kujala et al., 2004).

The primary goal of the current study was to provide clarification regarding the functional role of the PMN. However, a secondary goal was to establish an ERP paradigm that is sensitive to small units of phonological information, and specifically individual phonemes. Since measures of phoneme awareness, such as segmentation and phoneme deletion, are among the best predictors

---

[2] *This study was conducted in Finnish; English examples are provided for illustrative purposes.*

of reading skill (Hulme et al., 2002; Muter et al., 1998), establishing an ERP paradigm that is sensitive to phoneme awareness will have important implications for future studies of reading development and reading disability. In the current study, we manipulated the lexicality of stimuli while participants performed a phoneme deletion task modeled after that employed in Newman et al. (2003). Participants were instructed to delete the initial consonant from a CCVC word or nonword (e.g., *snap/snoth, / s/*), followed by presentation of the correct choice (e.g., *nap/noth*) or an incorrect choice that contained a mismatch at the initial phoneme (e.g., *tap/toth*). It was hypothesized that the PMN would be maximal to mismatch items. If the PMN represents the goodness-of-fit between phonological expectancies and speech input, as Connolly and colleagues suggest, then no differences in response characteristics (latency, amplitude) should be found for mismatching word and nonword trials. However, if the PMN is modulated by the activation of specific word candidates, then differences between word and nonword mismatch trials will be observed. Finally, if the N400 response reflects lexical selection mechanisms, then one would predict that the N400 response would be modulated by the lexical status of the speech input. That is, one would expect that mismatching words would elicit a larger N400 response than would mismatching nonwords.

2. Materials and Methods

2.1. Participants

Fourteen right-handed English-speaking participants (13 females; M age = 23.5 years [S.D. = 6.90; range = 19–39]) volunteered for this study. All participants reported normal hearing and were screened with a self-report health questionnaire for a history of neurological, audiological and/or psychological problems. Participants were screened for reading ability and phonological processing skills using the Word Attack and Word Identification subtests of the Woodcock Diagnostic Reading Battery (Woodcock, 1998), and the Elision subtest of the Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 1999). No participants were excluded on the basis of their performance on these tests. The average raw scores and standard deviation for each of the tests are as follows: Word Attack, M = 39.15, S.D. = 3.13; Word ID, M = 100.07, S.D. = 5.81; CTOPP Elision, M = 18.84, S.D. = 0.80. All participants provided informed consent, and the responsible ethics board approved the study.

2.2. Stimuli and procedures

One hundred and sixty monosyllabic CCVC items (80 words, 80 nonwords), were binaurally presented to participants along with the sound of the initial phoneme that was to be deleted (e.g. snap, /s/). A delay of 1 s separated the presentation of the prime item from that of the to be deleted phoneme. Following an interstimulus interval (ISI) of 2 s, participants heard a correct (e.g., nap) answer thatmatched, or an incorrect (e.g., tap) answer that mismatched, the anticipated answer after removing the initial phoneme. Responses always consisted of a CVC format, and participants were informed that the phoneme targeted for deletion was always the initial speech sound. The duration of word (942 ms) and nonword (971 ms) primes did not differ (p > .05), nor did the duration of word (794 ms) and nonword (828 ms) responses (p > .05). Phonetic properties of the stimuli were not explicitly controlled for in this study; however, the majority of primes and responses began with a voiced or voiceless stop consonant. For the 160 primes, there were 160 possible answers divided into four categories (i.e., 40 match word (MW) + 40 match nonword (MNW) + 40 mismatch word (MMW) + 40 mismatch nonword (MMNW). Items were differentially randomized for each participant. All stimuli were spoken by a female voice recorded with Cool Edit 2.0 program, digitized at 20,050 Hz and presented through headphones. EEG recording took place in a sound attenuated room. Participants were instructed to determine the resulting segment after the initial consonant had been removed. They were given a two-button response pad and instructed to press the right button to segments that matched their expectations and the left button for those that did not match their expectations. Button presses were counterbalanced across participants. In order to minimize artifacts associated with participants making their behavioral responses, participants were asked to withhold their response until they had heard the entire auditory stimulus. In addition, accuracy was emphasized over speed in order to reduce the number of trials that would have to be discarded as a result of behavioral errors.

2.3. Electrophysiological recording

EEG activity was recorded at 25 sites (F3, Fz, F4, F7, F8, FC3, FCz, FC4, FT7, FT8, C3, Cz, C4, CP3, CPz, CP4, P3, Pz, P4, TP7, TP8, T3, T4, T5, and T6) using sintered Ag/AgCl electrodes embedded in a NeuroScan Quick-Cap and referenced to linked ears according to the American Electroencephalographic Society (1991). [3] Impedances were kept at or below 5 kV. The electrooculogram (EOG) was recorded with electrodes placed above and below the right eye (vertical), and on the outer canthus of each eye (horizontal), and an electrode embedded in the Quick-cap (i.e., AFz) served as ground. EEG was amplified by a Sensorium EPA-5 amplifier and acquired and processed with BrainProducts' Brain Vision software. Stimuli and trigger codes were delivered by Neurobehavioral's Presentation program. EEG was recorded continuously with a bandpass of 0.1–100 Hz and sampled at 500 Hz. Data were epoched off-line from 100 to 800 ms post-stimulus onset, and digitally filtered with a low pass filter setting of 20 Hz. Trials contaminated with EOG activity greater than 75 mV were rejected from the analysis. EEG data on trials that participants responded to correctly were averaged separately across the four conditions. Fourteen percent of trials were excluded from further analyses due to artifacts and/or behavioral errors.

2.4. Data analysis

The amplitude of the PMN and N400 were measured by deriving the average integrated amplitude, relative to baseline, within specified time intervals for each component. Peak latency was defined as the time from stimulus onset to the peak amplitude of each response within the given time window. For both amplitude and latency measures, the PMN was scored between 260 and 320 ms, and the N400-like response between 380 and 460. These time intervals were chosen based on visual inspection of the grand averaged waveforms.

Repeated measures analysis of variance (ANOVA) using conservative degrees of freedom (Greenhouse and Geisser, 1959) was conducted for each component's latency and amplitude. Each ANOVA consisted of three factors, Lexicality (word, nonword), Congruency (match, mismatch), and Site (25 electrode locations). In situations where either the Lexicality or Congruency factor interacted with Site, a second analysis was conducted in order to further investigate the component's scalp topography. In this secondary analysis, the Site factor was divided into two factors: Region (frontal, central, temporal, parietal) and Hemisphere (left, right). Each Region x Hemisphere combination was linearly derived from a combination of two sites: left frontal (F3, FC3), right frontal (F4, FC4), left central (C3, CP3), right central (C4, CP4), left temporal (T7, FT7), right temporal (T8, FT8), left parietal (P3, P7), and right parietal (P4, P8). The midline sites were excluded from this analysis. Finally, post hoc comparisons were conducted using the Tukey honestly significant difference (HSD) test, with $p < .05$ as the required level of significance.

---

[3] *The use of linked electrodes as a reference is considered inadvisable as current shunting between electrodes may distort scalp voltage distributions. The distribution of the components under investigation in this study are well known and were replicated here. The other risk of this type of referencing is that source analyses may be inaccurate; this study did not have source localization of these components as its objective (see Picton et al., 2000).*

3. Results

3.1 *Behavioral results*

Accuracy and RT data were available for 13 of the 14 participants. Behavioral data from one participant was not available due to a technical problem. Lexicality (word, nonword) and

Congruency (match, mismatch) were entered as factors in a repeated measures ANOVA to analyze accuracy and RT data. No significant effects were observed for the analysis of accuracy data. Participants performed at near ceiling levels with an overall accuracy rate of 95.6%. Analysis of RT data indicated a significant main effect of Lexicality ($F(1,12) = 12.73$, $p < .01$, $e = 0.52$), as words (1154.09 ms) were responded to faster than nonwords (1229.77 ms). None of the other behavioral effects were significant. Note that participants were asked to withhold their behavioral response until complete presentation of the auditory stimulus; this request exaggerated the length of participants' response times.

3.2. *Electrophysiological findings*

Analysis of ERP data was based on 13 of 14 participants; it was decided to exclude the ERP data of the participant for whom behavioral data were lost due to technical problems. Fig. 1 illustrates the grand average waveforms of 13 participants for words and nonwords that matched or mismatched the anticipated answer that was primed by the prompt. Mismatch items (MM) evoked a fronto-central PMN response, which appeared not to differ between words and nonwords. In contrast, an N400-like response appeared to be differentially responsive to mismatched word (MMW) and mismatched nonword (MMNW) items over central and posterior sites.

3.3. *PMN*

Visual inspection of the grand average waveforms reveals that the MM conditions produced larger PMN responses than did the M conditions, regardless of the item's lexicality (Fig. 1). The ANOVA conducted on the PMN data revealed a significant effect of Congruency ($F(1,12) = 35.41$, $p < .001$, $e = 0.75$), reflecting the fact that PMN amplitudes in the MM conditions were more negative than responses elicited in the M conditions. The site effect was significant ($F(24,288) = 4.46$, $p < .05$, $e = 0.27$) with amplitudes tending to be more negative over frontal and central sites. Importantly, neither the main effect of Lexicality ($p = .88$), nor the interaction of Lexicality × Congruency was significant ($p = .62$), indicating that the PMN was not sensitive to the manipulation of lexicality. The Congruency × Site interaction was significant ($F(24,288) = 12.74$, $p < .001$, $e = 0.51$). This effect is more easily interpreted by conducting the secondary analysis involving regionalized data. A secondary analysis was conducted with Lexicality (2) × Congruency (2) × Region (4) × Hemisphere (2) as factors in order to examine differences in scalp topography. Interpretations of the following interactions are all supported by post hoc tests that were significant at the $p < .05$ level. As in the omnibus analysis, the Region/Hemisphere analysis revealed that the PMN was largest in MM conditions. The interaction of Congruency × Region was significant ($F(3,36) = 16.71$, $p < .001$, $e = 0.58$). While MM items were evenly distributed across all 4 regions, M items were more positive in the central region relative to the three other regions. The analysis of PMN latency revealed a significant effect of Congruency ($F(1,12) = 12.87$, $p < .003$, $e = 0.50$), with the PMN peaking later in the MM conditions (288 ms) than in the M conditions (277 ms). A significant effect of Site ($F(24,288) = 4.16$, $p < .01$, $e = 0.24$), was associated with the PMN peaking earlier over parietal locations. The latency of the PMN was not modulated by the Lexicality Factor ($p = .90$) or by the interaction of Congruency × Lexicality ($p = .64$).

3.4. N400

As depicted in Fig. 1, the PMN is followed by a negative-going wave peaking at approximately 420 ms that appears to be augmented in the MMW condition relative to the MMNW condition at central and parietal locations. An omnibus analysis of the 380–460 interval revealed a significant Site effect (F(24,288) = 4.36, p < .01, e = 0.27), a significant Congruency × Site interaction (F(24,288) = 7.47, p < .001, e = 0.38), and a significant three-way interaction of Congruency × Lexicality × Site (F(24,288) = 3.17, p < .05, e = 21). To better assess the scalp distribution of the N400, a secondary Regional analysis was performed. This analysis revealed a main effect of Region (F(3,36) = 4.66, p < .05, e = 0.28), which itself interacted significantly with both Congruency [Congruency × Region (F(3,36) = 9.57, p < .001, e = 0.44)] and Lexicality [Lexicality × Region (F(3,36) = 4.33, p < .05, e = 0.26)]. Finally, all three of these factors interacted significantly [Congruency × Lexicality × Region (F(3,36) = 3. 60, p < .05, e = 0.23)]. The three-way interaction was followed up by performing separate two-way ANOVAs with Congruency and Lexicality as factors for each of the four regions: Frontal, Central, Parietal and Temporal. No significant effects in N400 amplitude were observed for the Frontal or Temporal region; though the main effect of Congruency approached significance for the Temporal region (p = .08). The results for the both the central and parietal regions revealed a significant main effect of Congruency (central, F(1,12) = 6.57, p < .05, e = 0.35; parietal, F(1,12) = 6.86, p < .05, e = 0.36), which was attributable to more negative amplitudes recorded for mismatching compared to matching items. The Congruency × Lexicality interaction was significant for the central region (F(1,12) = 6.10, p < .05, e = 0.34) and approached significance for the parietal region (F(1,12) = 3.99, p = .07,e = 0.36). Post hoc analyses performed for N400 amplitudes recorded over the central region found that that the N400 was larger in the MMW condition compared to the MMNW condition over central and parietal regions.

Analysis of the peak latency of the N400 response revealed a significant main effect of Site (F(24,288) = 7.99, p < .01, e = 0.36). This was due to the response peaking earlier over parietal locations compared to frontal and central locations. A significant Lexicality × Congruency effect (F(1,12) = 7.65, p < .05, e = 0.25), was attributable to the response peaking earlier in the MMW condition compared to the other conditions.

3.5. Difference waves

In order to isolate the PMN and N400 and better understand their respective sensitivities to the experimental conditions, word and nonword difference waveforms were derived by subtracting the MW from the MMW condition, and the MNW from the MMNW condition. As illustrated in Fig. 2, the grand average difference waves do not distinguish words from nonwords over frontal and temporal regions. However, differences between words and nonwords begin to emerge over central and parietal sites at approximately 350 ms, likely reflecting the larger N400 recorded over these regions for mismatching words compared to mismatching nonwords.

In order to better characterize the time course of the PMN and N400, the mean amplitude for word and nonword items was calculated across five 50 ms time intervals: 200–250, 250–300, 300–350, 350–400, and 400–450. Note that an early time window was included in this analysis because differences between MMW and MMNW were apparent, particularly at Pz and over the right central region, in both Figs. 1 and 2. Separate three-way ANOVAs were conducted for each of these time intervals with Lexicality (word, nonword) × Region (frontal, central, temporal, parietal) × Hemisphere (right, left) as within-subjects factors. None of the effects for the 200–250 time window were significant. The results for the next three time intervals (250–300; 300–350; 350– 400) were generally the same; a significant main effect of Region was obtained (250–300, F(3,36) = 12.73, p < .001, e = 0.52; 300– 350, F(3,36) = 16.09, p < .001, e = 0.57; 350–

400, F(3,36) = 16.72, p < .001, e = 0.58) and none of the other effects, including the interactions, were found to be significant. In the case of the 250– 300 time interval the main effect of Region was attributable to more negative amplitudes recorded over the central region; however, for the later two time intervals (i.e., 300–350, 350–400), amplitudes were more negative over both central and parietal regions. Importantly, in the earliest time windows, there was no main effect of Lexicality (200–250, p = .57; 250–300, p = .38), nor was there a significant Lexicality × Region interaction (200–250, p = .25; 250–300, p = .63). The analysis for the 400–450 time interval found a significant main effect of Region (F(3,36) = 8.51, p < .01, e = 0.42), which was attributable to more negative amplitudes recorded over parietal regions. The interaction of Lexicality × Region was significant (F(3,36) = 3.71, p < .05,e = 0.24). Further analyses revealed that the word difference waves were more negative compared to the nonword differences waves over central and parietal regions. These results offer confirmatory evidence that while the PMN is not sensitive to lexicality, the N400 response is modulated by the lexical status of the mismatching item.

4. Discussion

The primary objective of this study was to clarify functional interpretations of the PMN. As outlined in Section 1, there is disagreement about whether the PMN reflects an autonomous stage of spoken word recognition, or whether it represents an early lexical selection process. Findings reported here show that the PMN is largest to items that mismatch expectations primed by the task context (i.e., the prompt) regardless of their lexicality status. These results are consistent with the findings of Connolly et al. (2001), who found that the response characteristics of the PMN did not differentiate between words and nonword targets that failed to match primed expectations. In the study reported here, there were no statistical differences in the amplitude and latency characteristics of the PMN, with the PMN peaking at approximately 288 and 287 ms for both words and nonwords, respectively. Visual inspection of the waveforms indicates that the PMN to words appears larger than that to nonwords over centro-parietal sites. However, this effect is likely due to the onset of the centroparietally distributed N400 response which was larger to words than nonwords and may have affected the manifestation of the PMN to some degree.

We liken the N400 response observed in the present study for mismatching words and nonwords to the N450 seen during visual rhyming tasks (Rugg, 1984a,b) and the N400-like response observed during auditory lexical decision studies (Praamstra et al., 1994; Praamstra and Stegeman, 1993). Our finding that the N400 response is larger to mismatching words compared to mismatching nonwords is consistent with Dumay et al. (2001), and supports their hypothesis that larger N400 priming effects for words are due to lexical selection mechanisms. Under this proposal words sharing phonological overlap with targets are initially activated in the lexicon and their subsequent deactivation is reflected in the increased amplitude of the N400. Pseudowords would not produce the same degree of lexical activation, and in turn would not show N400 effects.

On the basis of current results as well as that of our previous work, we propose that the PMN and N400 response represent two distinct mechanisms. We believe that the PMN reflects a phonological stage of word recognition at which a shortlist (Norris, 1994) or cohort (Marslen-Wilson, 1987) of phonologically specified candidates formed by task expectations interacts with the incoming acoustic–phonetic information present in the speech signal. If the incoming information matches expectancies, then the PMN is attenuated. If, on the other hand, expectations are not met, then the PMN is augmented and a new set of candidates must be established. This goodness-of-fit measure is not, however, influenced by the activation of lexical competitors, such as rhyming items (Newman et al., 2003) or semantically related items (Connolly and Phillips, 1994), and as seen in the current study, is not influenced by the lexical status of the target (also Connolly et al., 2001). If the PMN represented a point of lexical selection as previously proposed (Hagoort and Brown, 2000), then one would anticipate differences in response characteristics between words and nonwords. No such differences have been observed in this study, nor have they been found to occur in previous work (Connolly et al., 2001; Kujala et al., 2004). Furthermore, the existence of the PMN to nonwords suggests that the list of potential candidates (i.e., anticipated response) was formed by phonological expectations rather than by lexical expectations (Connolly et al., 2001).

Like the PMN, the N400 response also appears to reflect a goodness-of-fit measure. Unlike the PMN, however, the N400 response may reflect a stage of word recognition that allows for the activation of word candidates that do not begin with the same word onset and yet share phonological overlap with a target response. It is tempting to align the N400 response with predictions made by TRACE; that the set of activated candidates is in part determined by the overall match between speech input and potential lexical representations. That is, the amplitude of the N400 response would be expected to increase as the degree of overlap between speech input and a lexical representation decreased. The current study did not include a manipulation of the overall match between the anticipated response and the actual response that was presented, and so our findings cannot speak directly to this prediction. However, numerous studies have observed a reduction in the N400 for rhyming words compared to non-rhyming words, supporting the proposal that the N400 observed here reflects phonological word form-overlap (Dumay et al., 2001; Praamstra et al., 2004).

The distinct processes reflected by the PMN and the N400 may be incorporated into continuous mapping models of spoken word recognition, such as TRACE. According to TRACE, the initial speech input plays a significant role in determining which words become activated. However, word-initial information is not emphasized to the same degree as in the cohort model. In TRACE, words that do not have the same onset, but that partially overlap with the spoken word (e.g., rhyming words) will also become weakly activated as speech unfolds. The presence of such a “recovery mechanism” (Allopenna et al., 1998, p. 420) improves the listeners’ ability, for instance, to cope with noisy speech environments or misarticulated speech. Within the context of TRACE, we propose that the PMN reflects a mechanism that is highly intolerant to phonological mismatches between initial speech information and potential target representations, whereas the N400 represents a more error tolerant mechanism that is more sensitive to the overall overlap between speech information and potential lexical candidates. These mechanisms working together provide an ‘alerting’ response to the speech input system indicating an early violation of phonologically based contextual expectations (the PMN) followed by a ‘recovery’ response (the N400) reflecting a top–down process that uses lexical information to reanalyze the input.

The phoneme deletion paradigm employed here may be more sensitive to the mechanisms underlying the PMN, insofar as the response itself appears more robust than that seen in previous studies, and was minimally influenced by the later peaking N400 response. The phoneme deletion task involves explicit awareness of individual phonemes, an awareness that developmentally speaking, is more difficult to attain than rime or syllable awareness (Anthony et al., 2003) and has been shown to be among the best predictors of reading success (Hulme et al., 2002). The finding that an ERP component, the PMN, appears sensitive to phoneme awareness will have important implications for future research.

The increased demands on working memory during the phoneme deletion task might explain why the PMN is particularly robust in this study compared to previous work. Previous work has shown that working memory demands modulate the PMN (D’Arcy et al., 2004). In that study, participants were presented with a visual sentence (e.g., The man is in the classroom), and were instructed to anticipate the superordinate (i.e., a word higher in the semantic hierarchy than the prime, classroom) terminal word (e.g., school). Following the prime, a spoken sentence was presented that ended with a target word that was congruent to the prime or that was incongruent to the prime (e.g., The man is in the school/barn). Target words in the incongruent condition mismatched initial phonological expectations as well as semantic expectations. In order to measure the influence of active lexical candidate numbers on the PMN and N400, the probability of target stimuli were divided into high/low congruent conditions, such that the Low Congruent condition was characterized as having more possible lexical candidates than the High Congruent condition. For example, in the High Congruent condition, target sentences ended with a high probability congruent ending (e.g., The boy is swimming in the shallow end. The boy is in the pool). In the Low Congruent condition, the target sentences ended with a low probability congruent ending (e.g., The woman is swimming in the sunken ship. The woman is in the ocean [water, sea, lake also being possible candidates]. As expected, incongruent terminal words elicited both PMN and N400 responses. The manipulation of high/ low probability, however, differentially influenced the PMN and N400. In the Low Congruent condition, a PMN, but no N400 was observed. On the basis of these findings, D'Arcy et al. (2004) suggested that the PMN is modulated by the number of activated candidates, possibly due to increased demands on the maintenance and rehearsal of those candidates in phonological working memory. Adding further support to their proposal that working memory processes underlie the elicitation of the PMN, D'Arcy et al. (2004) found that the primary PMN sources were localized to regions previously shown to subserve phonological working memory processes (i.e., the inferior frontal and inferior parietal lobes).

4.1. Summary

The findings of this experiment offer support for the proposal that the PMN is not dependent on lexical-semantic mechanisms. These findings parallel those of Connolly and colleagues (Connolly and Phillips, 1994; Connolly et al., 2001; Newman et al., 2003), and argue against interpretations of the PMN as representing early lexical selection mechanisms (van den Brink et al., 2001; Hagoort and Brown, 2000). Rather, the findings presented here suggest that the PMN is a neural measure operating at a phonological stage of spoken word recognition that precedes lexical selection and is highly tuned to the onset of a spoken word.

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.

American Electroencephalographic Society, 1991. Guidelines for standard electrode position nomenclature. Journal of Clinical Neurophysiology 8, 200–202.

Anthony, J.L., Lonigan, C.J., Driscoll, K., Phillips, B.M., Burgess, S.R., 2003. Phonological sensitivity: a quasi-parallel progression of word structure units and cognitive operations. Reading Research Quarterly 38, 470–487.

Bentin, S., Kutas, M., Hillyard, S.A., 1993. Electrophysiological evidence for task effects on semantic priming in auditory word processing. Psychophysiology 30, 161–169.

Connine, Blasko, D.G., Titone, D., 1993. So the beginnings of spoken words have a special status in auditory word recognition. Journal of Memory and Language 32, 193–210.

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., Forbes, K.A.K., 1995. The effects of phonological and semantic features of sentence-ending words on visual event-related brain potentials. Electroencephalography and Clinical Neurophysiology 94, 276–287.

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. NeuroReport 12 237–243, 40–51.

Connolly, J.F., Stewart, S.H., Phillips, N.A., 1990. The effects of processing requirements on neuropsychological responses to spoken sentences. Brain and Language 39, 302–318.

D'Arcy, R.C., Connolly, J.F., Service, E., Hawco, C.S., Houlihan, M.E., 2004. Seperating phonological and semantic processing in auditory sentence processing: a highresolution event-related brain potential study. Human Brain Mapping 22, 40– 51.

Dumay, N., Benraı¨ss, A., Barriol, B., Ce´ cile, C., Radeau, M., Besson, M., 2001. Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. Journal of Cognitive Neuroscience 13, 121–143.

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

Hagoort, P., Brown, C.N., 2000. ERP effects of listening to speech: semantic ERP effects. Neuropsychologia 38, 1518–1530.

Hulme, C., Hatcher, P.J., Nation, K., Brown, A., Adams, J., Stuart, G., 2002. Phoneme awareness is a better predictor of early reading skill than onset-time awareness. Journal of Experimental Child Psychology 82, 2–28.

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 (PMN) response with MEG. Cognitive Brain Research 21, 106–113.

Kutas, M., Hillyard, S., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.

Kutas, M., Van Petten, C., 1988. Event-related brain potential studies of language. In: Ackles, P.K., Jennings, J.R., Coles, M.G.H. (Eds.), Advances in Psychophysiology. JAI Press, Greenwich, CT.

Marslen-Wilson, W.D., 1987. Functional parallelism in spoken word recognition. Cognition 25, 71–102.

Marslen-Wilson, W.D., 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.

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

Muter, V., Hulme, C., Snowling, M., Taylor, S., 1998. Segmentation, not rhyming, predicts early progress in learning to read. Journal of Experimental Child Psychology 71, 3–27.

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 potentials. Psychophysiology 40, 640–647. 120 R.L. Newman, J.F. Connolly / Biological Psychology 80 (2009) 114–121

Norris, D., 1994. Shortlist: a connectionist model of continuous speech recognition. Cognition 52, 189–234.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson Jr., R., Miller, G.A., Ritter, W., Ruchkin, D.S., Rugg, M.D., Taylor, M.J., 2000. Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. Psychophysiology 37, 127–152.

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 (3), 204–219.

Praamstra, P., Stegeman, D.F., 1993. Phonological effects on the auditory N400. Cognitive Brain Research 1, 73–86.

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

Rugg, M.D., 1984a. Event-related brain potentials and the phonological processing of words and nonwords. Neuropsychologia 22, 642–647.

Rugg, M.D., 1984b. Event-related brain potentials in phonological processing tasks. Brain and Language 23, 225–240.

van den Brink, D., Brown, C., 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.

Wagner, R.K., Torgesen, J.K., Rashotte, C.A., 1999. Comprehensive Test of Phonological Processing (CTOPP). PRO-ED, Austin, TX.

Woodcock, R.W., 1998. Woodcock Diagnostic Reading Battery. Riverside Publishing, Itasca, IL

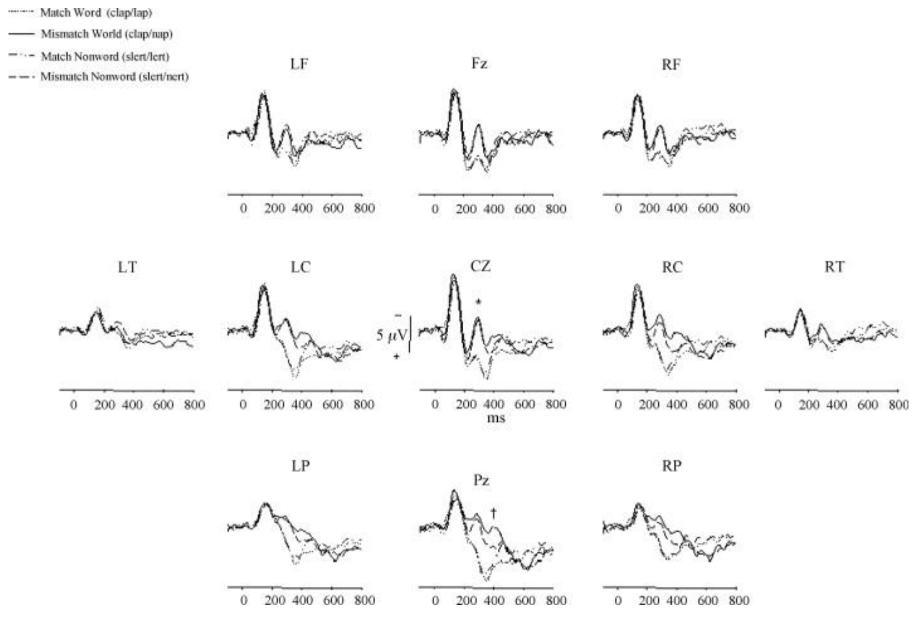

Fig. 1. Grand average ERP (*N* = 13) to target words/nonwords in the match word (MW, dotted line), mismatch word (MMW, solid line), match nonword (MNW, dashed dot line), and mismatch nonword (MMNW, dashed line) conditions for three midline sites (Fz, Cz, Pz), and eight regions [left and right frontal (LF, RF), central (LC, RC), temporal (LT, RT), and parietal (LP, RP) regions]. The PMN (*) was largest to MM items, and did not differentiate mismatching words from nonwords. An N400-like response (†) differentiated MMW from MMNW. Time (ms) is on the *x*-axis and amplitude (μV) is on the *y*-axis. Negative is up.

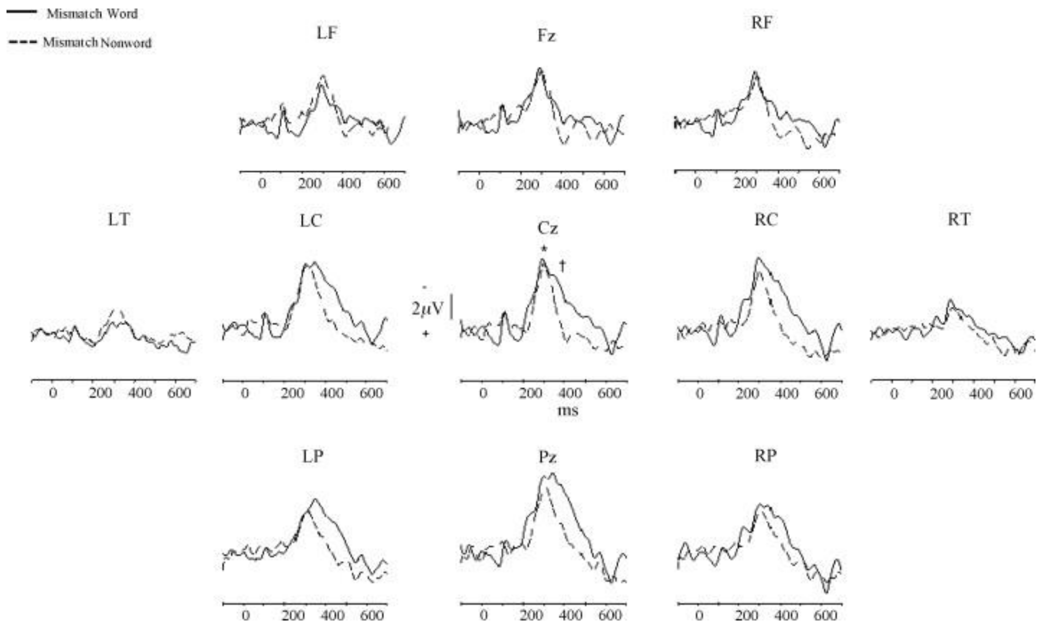

Fig. 2. Grand average (*N* = 13) difference waveforms for word (solid line) and nonword (dashed line) items at left and right frontal (LF, RF), left and right central (LC, RC), left and right temporal (LT, RT) and left and right posterior (LP, RP) regions. Word difference waves were derived by subtracting the MW from the MMW condition, and nonword difference waves were derived by subtracting the MNW from the MMNW condition. Time (ms) is on the *x*-axis and amplitude (in μV) is on the *y*-axis. Negative is up.