ELSEVIER

Biological Psychology 60 (2002) 121–150

BIOLOGICAL
PSYCHOLOGY

www.elsevier.com/locate/biopsycho

Electrophysiological evidence for the efficiency of spoken word processing

Timothy B. O'Rourke, Phillip J. Holcomb *

Tufts University, 490 Boston Avenue, Medford, MA 02155, USA

Received 15 March 2002; accepted 3 June 2002

Abstract

The Cohort model (Marslen-Wilson, Spoken Word Recognition, MIT Press, Cambridge, MA, 1987, pp. 71-103) proposes that spoken words are frequently recognized at the first point in the acoustic signal where a unique lexical representation is specified. This proposal was tested in two experiments. In experiment 1 participants made speeded lexical decisions to spoken words and pseudowords. In experiment 2 participants passively listened without making overt responses. In both experiments the recognition points for words (the point past which no other lexical item was consistent with the acoustic signal) and deviation points for pseudowords (the point past which no real word is compatible with the acoustic signal) were manipulated. An ERP negativity in the region of the N400 component and RT occurred sooner for items with early than late recognition/deviation points when measures were time-locked to stimulus onset. In experiment 1, when time-locking was to recognition/deviation points, early and late words produced N400s and RTs with indistinguishable latencies, while late pseudowords produced faster RTs and earlier N400s than early pseudowords. Experiment 2 replicated the N400 effects for words, but only produced a trend in the same direction as experiment 1 for pseudowords. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Event-related potentials; Cohort model; N400; Spoken word recognition

1. Introduction

A fundamental aspect of spoken language is that it unfolds over time. Unlike reading, where an entire word can frequently be encoded in a single fixation, spoken

* Corresponding author. Tel.: +1-617-627-2319; fax: +1-617-627-3181
E-mail address: pholcomb@tufts.edu (P.J. Holcomb).

PII: S0301-0511(02)00045-5

words usually require several hundreds of milliseconds to be heard. Therefore, an important question for researchers interested in understanding the dynamics of speech comprehension has been at what point in time are spoken words recognized (Tyler and Frauenfelder, 1987)? One possibility is that listeners simply wait for all of the acoustic information in a word to be presented and only then engage recognition processes. However, this type of 'wait and see' model seems implausible for a number of reasons, not the least of which is that natural connected speech has few pauses between words, thus leaving insufficient time for listeners to recognize a word before the next one begins. It seems more likely that the demands of natural speech require spoken word recognition to be more efficient.

One theory of word recognition that has argued for efficiency is Marslen-Wilson's Cohort model (Marslen-Wilson, 1987, 1990, 1993; Marslen-Wilson and Welsh, 1978; Marslen-Wilson and Tyler, 1980). It was designed to account for the unique temporal properties of spoken language and to make explicit predictions about the nature and timing of word recognition. The model assumes that each representation in the lexicon has pattern matching abilities and that bottom-up information is simultaneously relayed, in parallel, to all elements. When a match occurs between incoming acoustic-phonetic information and a lexical element the element becomes active. Early on in a spoken word (e.g. after the first phoneme has been uttered) the number of lexical items activated will correspond to the number of words in the language that are compatible with the initial acoustic-phonetic input (the so-called word-initial cohort). As more acoustic-phonetic information enters the system, lexical elements which are no longer compatible with the acoustic input are dropped from the word-initial cohort. An item is said to be selected for recognition when the incoming acoustic-phonetic information is compatible with only one member of the initial cohort. At this point stored information about the selected lexical element (e.g. its meaning) is integrated into the current discourse representation.

A crucial element of the Cohort model is that processing is done with optimal efficiency, which means that selection can often occur before the end of a word. This is possible because many words are acoustically 'unique' prior to their final sounds. Optimal discrimination points, referred to here as recognition points, are therefore, defined as the earliest point in the acoustic signal when a word can be distinguished from all other words that begin with the same sound sequence.[1]

[1] Various authors have used the terms uniqueness point and recognition point to the earliest point at which a word can be identified based on its acoustic properties. Which term is used depends on the method used to define the point of optimal efficiency. Uniqueness point refers to the first phoneme in a word that differentiates it from all other words. This point is typically determined by referring to a phonemic dictionary. Recognition points are usually determined empirically by having participants guess the identity of a word based on successively longer acoustic segments. The point past which a word is identified with a relatively high rate of confidence is referred to as the recognition point. The current study used this latter approach for determining optimal efficiency points.

Several studies suggest that spoken word recognition operates in a manner consistent with the predictions of optimal efficiency. For example, in an experiment using a pseudoword monitoring task Marslen-Wilson (1984) had participants listen to speech sounds which became pseudowords at points between the first and forth phoneme. Marslen-Wilson argued that if lexical processing occurred with optimal efficiency, then these pseudowords should be recognized as such at these deviation points. This prediction was supported by the finding that non-word RT was a constant latency (450 ms) after deviation points regardless of whether this point was early or late in the acoustic stream (see also Radeau et al., 1989 and Radeau and Morais, 1990, for comparable studies using word stimuli).

Evidence inconsistent with the Cohort model's notion of optimally efficiency has also been reported. Taft and Hambly (1986) compared short pseudowords which had deviations points on the last phoneme with longer pseudowords that shared the same initial segment (e.g. MEP/MEPSIG). RTs were slower for long pseudowords which is inconsistent with the Cohort model's prediction of a constant RT for short and long items. In another study Goodman and Huttenlocher (1988) used pseudowords that were more like real words than those used by Marslen-Wilson. The deviation points occurred either early or late in the acoustic signal. Across four experiments RTs from stimulus onset followed the predictions of the Cohort model with faster RTs for early than late deviation point items. However, when time-locked to deviation points, RTs were not constant as predicted by the Cohort model. In fact, RTs to early deviation point pseudowords were actually longer than RTs to late deviation point items.

In summary, while there is modest evidence suggesting that listeners take uniqueness points into account, this is qualified by the finding that listeners do, at least on occasion, use information after this point when processing certain types of pseudowords. Note however, that all of the evidence going against the optimal efficiency prediction of the Cohort model was collected using the lexical decision task, while many of the studies supporting optimal efficiency used other tasks (e.g. naming and gender classification—Radeau et al., 1989 and Radeau and Morais, 1990). This suggests that one possibility for the seeming contradiction in results is the presence of idiosyncratic factors associated with specific tasks. The goal of the current study was to evaluate this possibility by testing words and pseudowords from both stimulus onset as well as from uniqueness/deviation points using a technique that is less susceptible to idiosyncratic task factors.

1.1. Event-related potentials (ERPs)

The findings from a number of studies suggest that the N400 component is sensitive to the time-course of certain aspects of word processing. Kutas and Hillyard (1980) were the first to demonstrate that this component is relatively larger whenever a word is semantically anomalous with respect to prior sentence context compared to when a word fits the prior context. Subsequent studies have shown that most words elicit some amount of N400 activity (e.g. Kutas and Hillyard, 1984).

Although there have been a number of interpretations of the functional significance of the N400, the most widely cited account is that it reflects the amount of effort involved in integrating semantic information into a higher order text or discourse representation (e.g. Brown and Hagoort, 1993; Holcomb, 1993).

Of particular relevance to the current research is a growing body of evidence linking the N400 to processes involved in recognizing isolated words or pairs of words (e.g. Bentin et al., 1985; Holcomb, 1988; Holcomb et al., in press; Rugg, 1987). For example, a number of studies have used the semantic priming lexical decision task (Meyer and Schvaneveldt, 1971) to show that the N400 is larger to visually presented targets words preceded by an unrelated priming word (e.g. doctor–table) than to target words preceded by a related prime (e.g. doctor–nurse). More recently a similar pattern of effects has been reported for spoken words (e.g. Bentin et al., 1993; Holcomb and Neville, 1990; Soares et al., 1991; Woodward et al., 1990). In one such study, Holcomb and Neville (1990) observed that the spoken word N400 effect (the ERP difference between related and unrelated word ERPs) started earlier than the offset of the shortest spoken stimuli, at a point before participants had heard the final sounds of even the briefest of words. This suggests that the N400 may be a useful indicator of the time-course of spoken word processing.

Three studies have used ERPs to examine issues associated with recognition points in words. Woodward et al. (1990), recorded ERPs to words in a memorization task. They reported that the latency of the N2 (mean latency 480 ms) was correlated with word duration and word recognition points. Soares et al. (1991) recorded ERPs while participants listened to French words and nonwords and made speeded lexical decisions. An items analysis found that recognition points in words and deviation points in pseudowords were, on average, only 23 ms different than the peak of a late negative wave (what they referred to as an 'N4') when ERPs were time-locked to these same points. However, the correlation between N4 latency and recognition point latency was not significant. While both of these studies suffer from a number of methodological problems (e.g. too few participants and items) they both suggest the intriguing possibility that spoken word N400s may be sensitive to the temporal properties of word processing hypothesized by the Cohort model to be important during word recognition.

More recently, Van Petten et al. (1999) recorded ERPs to sentence final anomalous and congruent spoken words. They reported that the N400 started 200 ms prior to the isolation points of words (the minimum duration required to identify a word in a gatting task), which suggests that at least some words were recognized and integrated prior to there having been enough acoustic information available to uniquely identify them. This result suggests that context can serve to facilitate spoken word processing beyond what is possible when words are processed in isolation. However, the Van Petten et al. study tested what might be argued to be the most extreme of circumstances (words occurring at the end of highly constraining sentences), and therefore most likely reflects word recognition processes at one end of a continuum. What remains unclear is what processes operate at lesser levels of contextual constraint or in the absence of context.

2. Experiment 1

In experiment 1 electrophysiological and behavioral measures were used to test the optimal efficiency prediction of the Cohort model when words are encountered without a constraining context. Participants made speeded lexical decisions to spoken words and pseudowords while ERPs were recorded from 13 scalp sites. Recognition points in words were operationally defined as the mean latency at which words were identified with 80 percent confidence in a gating task. Pseudowords were constructed by rearranging the phonemes of words such that each item clearly became a pseudoword at a specific latency (i.e. the deviation point). Based on recognition and deviation points, items were then divided at the median making four groups of stimuli: words with early recognition points (e.g. pupil), words with late recognition points (e.g. carriage), pseudowords with early deviation points (e.g. boursley) and pseudowords with late deviation points (e.g. barble). RTs and ERPs were measured from stimulus onset and from the onsets of recognition/deviation points in these four conditions.

There were two predictions which follow from the Cohort Model. First, when time-locked to stimulus onset, the time-course of the N400 and RT latency will be shorter for words and pseudowords with early as compared to late recognition and deviation points. Second, when time-locked to recognition points, RT latency and N400 time-course will be similar for words and pseudowords with early and late recognition/deviation points. Note that this pattern of results would be extremely damaging for alternative accounts of spoken word processing which emphasize the role of post-recognition point information (Goodman and Huttenlocher, 1988; Taft and Hambly, 1986). An alternative pattern of predictions is made by post-recognition point models. According to these accounts RT and N400 time-course should be delayed for early as compared to late words and pseudowords when time-locked to recognition/deviation points. This is because it is not possible to make word/nonword decisions until after the final sounds of an item are processed as any potential item might become a pseudoword as late as the last phoneme. Since time-locking to recognition/deviation points results in a longer duration to stimulus offset for early than late items, processing of early items from this point should take longer. Such a finding would cast serious doubt on the Cohort theory's claim of optimal efficiency.

There is, however, another possible outcome. If the RT effect in the lexical decision task reflects both word recognition as well as post-lexical strategic processes associated with decision making (e.g. Chumbley and Balota, 1984) while the N400 reflects primarily word-based processing (e.g. Kounios and Holcomb, 1992) then we might expect to see a dissociation between the RT and N400 results. In other words, N400 time-course might follow the predictions of the Cohort model, while RT proves more consistent with post-recognition model predictions.

2.1. Methods

2.1.1. Participants

Thirty righted-handed native English speaking Tufts University undergraduates (18 female, 12 male) participated for course credit. All were between 18 and 22 years of age (mean = 20.2 years) and were self reported to have normal hearing.

2.1.2. Materials

All stimuli were bisyllabic and had stress on the first syllable. The durations of all the stimuli were between 600 and 900 ms (mean for words 708 ms and for pseudowords 748 ms) and had frequencies between 1 and 50 per million (Francis and Kucera, 1982) and familiarity ratings of 6.0 or greater in the Hoosier Mental Lexicon (Nusbaum et al., 1984).

From the 20,000 entries in the Hoosier Mental Lexicon (Nusbaum et al., 1984), 209 bisyllabic words with stress on the first syllable were chosen. Recognition points were determined for these items using a gating task (Grosjean, 1980). Each word was presented repeatedly in fragments, starting with the first 50 ms of the word and with the addition of 50 ms for each subsequent presentation, until the entire word had been presented. Participants ($N = 8$, separate from the sample used in the ERP experiment) were required to verbally identify each word, or what the word would become, at each gate and provide a confidence rating (1 = not confident, 10 = very confident). Recognition points for each word were defined as the mean duration of gates (across the eight participants) required for correct identification at a confidence rating of eight (Grosjean, 1980, 1985). Of 209 items tested, 167 had durations in the desired range (between 300 and 900 ms). Phonemes from 120 of these words were

Table 1
Example of scheme used to create pseudowords

Words
1. motIv (motive)
2. tRb|n (turban)
3. b@ lxd (ballad)
Set 1 Pseudowords
1. b@ lxn – from **b@ lx**d and tRb|**n**
2. motId — from **motI**v and b@ lx**d**
3. tRb|v — from **tRb**|n and motI**v**
Set 2 Pseudowords
4. tRlxd — from **tR**b|n and b@ **lxd**
5. b@ tIv — from **b@** lxd and mo**tIv**
6. mobdn — from **mo**tIv and tR**b|n**

rearranged to make 240 pseudowords (see Table 1).[2] For all pseudowords the first phoneme after the deviation point was the actual point where the item became a pseudoword (i.e. deviated from all real words). This was confirmed by consulting the Hoosier Mental Lexicon (Nusbaum et al., 1984). Of the remaining 47 words, 40 were used as filler words and seven were used to construct practice list items.

Three lists were constructed containing 80 words and 80 pseudowords. Each list contained 40 of the words used to create the pseudowords and the 40 filler words. Words used to create pseudowords as well as other pseudowords made from the same words were presented in different lists so that each participant heard all of the phoneme segments from the original list of 120 words, but no participant heard any segment more than once. Lists were counterbalanced between participants so that across participants all segments were heard an equal number of times in all conditions.

A male American English speaker familiar with the rules of English phonology sat in a sound attenuated booth and spoke all stimuli (words and pseudowords). The stimuli were recorded on an analogue tape and were subsequently digitized (24 kHz sampling rate 12-bit resolution). A temporal display of each digitized word and pseudoword was used displayed to establish stimulus onset, onset of recognition pointer deviation point, and stimulus offset. Stimuli were stored on the hard disk of a PC computer for subsequent real time presentation.

Two conditions, early and late, were derived by dividing the distribution of latencies from stimulus onset to the onset of recognition points and deviation points at the 50th percentile. Recognition points and deviation points from stimulus onset ranged from 116 to 437 ms for early pseudowords (mean = 346), from 395 to 662 ms for late pseudowords (mean = 508), from 325 to 500 ms for early words (mean = 427) and from 462 to 648 ms for the late word condition (mean = 533). Note that the overlap between early and late was due to slightly different median points for each of the three stimulus lists. Approximately half (52%) of the early pseudowords had deviation points one or two phonemes prior to the recognition point for the words from which they were derived and the other half had deviation points that coincided with the recognition point. Almost all (82%) of the late pseudowords had deviation points at the recognition points for the words from which they were derived (18% had deviation point one or two phonemes earlier than the recognition point).[3]

2.1.3. Procedure

Tin electrodes were held in place on the scalp with an elastic cap (Electrode-Cap International). Scalp locations included standard International 10–20 system

[2] Of the pseudowords 35% had deviation points that occurred one or two phonemes prior to the recognition point and the remaining 65% had deviation points corresponding to the recognition points of the word stem from which they were derived.

[3] In subsequent analyses pseudowords with early deviation points were sub-divided into those derived from words before and after their recognition points. No measure differentiated these items and all of the effects noted in the text for early and late deviation point pseudowords held up when only the post-recognition point items were included in the analyses.

locations over the left and right hemispheres at frontal (F7 and F8) and occipital sites (O1 and O2) and three locations on the midline at frontal (Fz), central (Cz), and Parietal (Pz). In addition, three pairs of electrodes were placed at the following non-standard locations because they have been shown to be sensitive to language manipulations and have shown consistent differences in previous ERP language studies (Holcomb and Neville, 1990; Holcomb, 1993): left and right temporo-parietal cortex (Wernicke's area) and its right hemisphere homologue (WL and WR: 30% of the interaural distance lateral to a point 13% of the nasion-inion distance posterior to Cz), left and right temporal cortex (TL and TR: 33% of the interaural distance lateral to Cz), and left and right anterior temporal cortex (ATL and ATR: 50% of the distance between T3/4 and F7/8). Eye blinks were monitored with an electrode placed below the left eye. Horizontal eye movements were monitored with an electrode lateral to the right eye. All electrodes were referenced to the left mastoid (the right mastoid was recorded actively to determine if there were any asymmetries between the mastoids). All electrode impedances were less than 5 K Ω. The electroencephalogram (EEG) was amplified by a Grass Model 12 amplifier system using a bandpass of.01 to 30 Hz (3 dB cutoff) and a sampling rate of 200 Hz.

Participants sat in a sound attenuated booth and made speeded lexical decisions to stimuli presented at a comfortable listening level (Sony headphones, model MDR-S30). A trial started with a fixation point presented in the center of a computer screen. One second after the onset of the fixation point a spoken stimulus was presented. The fixation point remained on throughout the spoken stimulus and was terminated after 2.5 s. Participants were told they should not blink while the fixation point was on the screen. Following the fixation point there was a 3.5-s blank screen inter-trial interval. Participants were told they could blink during this interval. The YES response hand was counter-balanced across participants and speed and accuracy were stressed equally. Participants were given a break approximately every 40 trials.

2.1.4. Data analysis

RTs were analyzed in two separate 2×2 repeated-measures analyses of variance (ANOVA). The first was conducted on RTs measured from stimulus onset for both early and late words and pseudowords. The second was conducted on RTs measured from the recognition points and deviation points of words and pseudowords.

Average ERPs were formed from trials in which a correct response occurred and which were free from ocular and movement artifacts. Separate ERPs time-locked to stimulus onset and time-locked to recognition points and deviation points were generated for early words, late words, early pseudowords and late pseudowords. The baseline used for stimulus onset averaging was between −100 and 0 ms and that used for recognition point/deviation point averaging was −650 to −550 ms. ERP waveforms were quantified in two ways. First, peak latency analyses were conducted on ERPs time-locked to stimulus onset and on a second set of ERPs time-locked to recognition/deviation points. The time window between which a peak was sought was 300–750 ms for ERPs time-locked to stimulus onset and −200 to +250 ms for ERPs time-locked to recognition/deviation points. Separate analyses were conduced for midline and lateral electrode sites. The midline analyses included three factors; stimulus type (words and pseudowords), latency of recognition point/deviation point (early and late), and electrode site (frontal; Fz, central; Cz, and parietal; Pz). The lateral peak latency analyses used different electrode sites (Frontal, Anterior Temporal, Temporal, Wernicke's and Occipital) and added the factor of hemisphere (left and right).

ERPs were also quantified using point-by-point t-test analyses (see Holcomb and Neville, 1991) in order to evaluate the earliest place in the ERPs where differences between early and late conditions emerged (i.e. the onset of an effect) as well as the point in time where differences ceased to be significant (i.e. offset effects). Because of the large number of sequential comparisons, only runs with more than five consecutive p values less than 0.05 were deemed noteworthy (Guthrie and Buchwald, 1991).

2.2. Results

2.2.1. Behavioral findings

Table 2 presents RTs and error rates for words and pseudowords in the early and late conditions measured from stimulus onset and from recognition/deviation points. The analysis for RTs measured from stimulus onset revealed a main effect for stimulus type ($F[1,29] = 124.20$, $p < 0.0001$), a main effect of latency to recognition/deviation point ($F[1,29] = 468.87$, $p < 0.0001$), but no significant interaction ($p = 0.25$). Errors also showed a significant main effect of stimulus type ($F[1,29] = 13.54$, $p < 0.001$). Participants responded faster and made fewer errors to words than to pseudowords. More importantly, stimuli with early recognition/deviation points were responded to significantly faster than those with late recognition/deviation points.

The analysis for RTs measured from the recognition/deviation points also produced main effects of stimulus type ($F[1,29] = 265.58$, $p < 0.0001$) and latency

Table 2
Mean RT (S.D.) in ms and percent errors (S.D.)

	From onset		From uniqueness point	
	RT	Errors	RT	Errors
Words				
Early	908 (130)	4.3% (4.8)	482 (132)	4.1% (4.5)
Late	1022 (137)	3.1% (6.0)	490 (135)	3.0% (6.1)
	From onset		From deviation point	
	RT	Errors	RT	Errors
Pseudowords				
Early	1069 (141)	8.9% (7.8)	723 (142)	8.9% (7.8)
Late	1169 (140)	8.5% (7.5)	668 (150)	7.7% (7.3)

of recognition/deviation point ($F[1,29] = 23.80$, $p < 0.0001$) as well as a significant interaction ($F[1,29] = 26.84$, $p < 0.0001$). Pair-wise comparisons revealed no RT difference between early and late words ($F < 1$), but early pseudowords produced significantly longer RTs than late pseudowords ($F[1,29] = 53.00$, $p < 0.0001$).

2.2.2. *ERP findings*

2.2.2.1. Peak latency analysis from stimulus onset. Less than 20% of trials were rejected for each condition because of artifact or incorrect responses (mean number of trials in each average per participant; early words = 34, late words = 36, early pseudowords = 34, late pseudowords = 34).

Figs. 1 and 2 display grand average ERPs for words and pseudowords with early and late recognition/deviation points time-locked to item onset. The waveforms include an initial central to anterior negativity with a peak latency near 100 ms (N1) followed by a positivity peaking near 200 ms (P2), a large broad negativity (N400) and at some posterior sites (e.g. Pz) a late positivity (P3). One important feature to

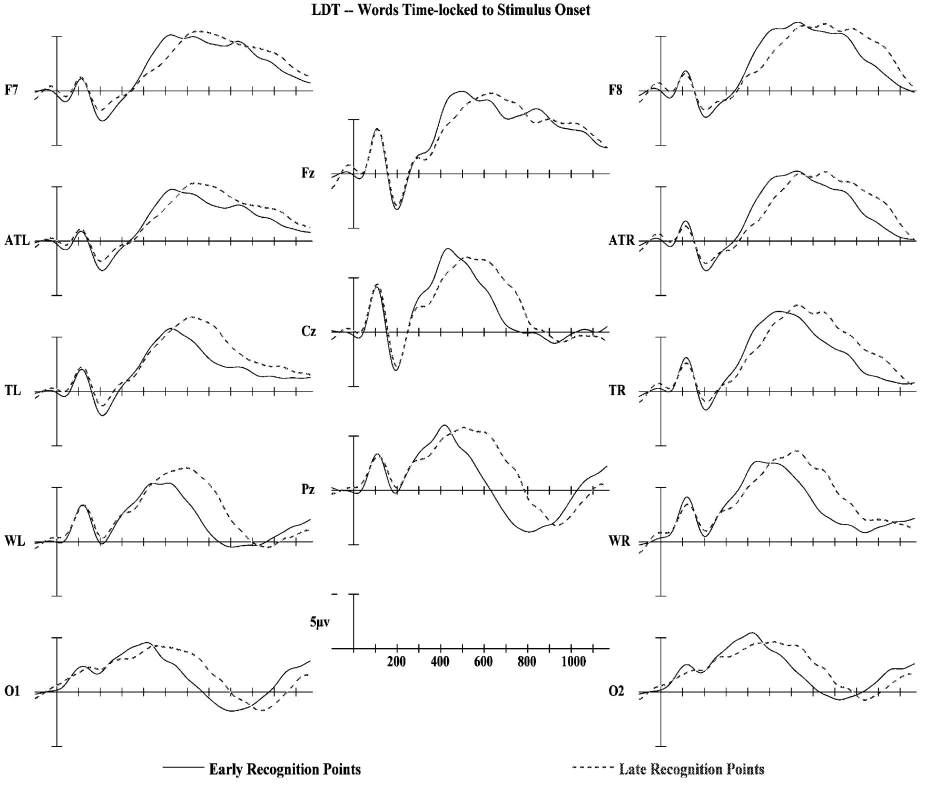

Fig. 1. Experiment 1 ERPs time-locked to stimulus onset for words with early (solid) and late (dotted) uniqueness points. Stimulus onset is marked by the vertical calibration bar and negative polarity is plotted upward.

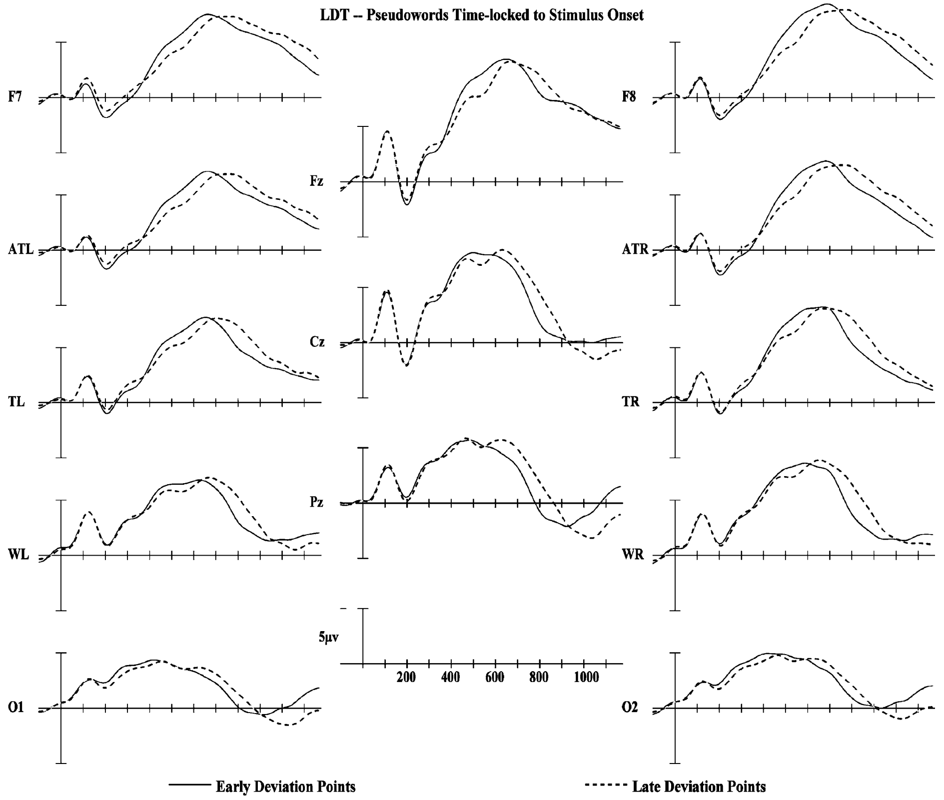

Fig. 2. Experiment 1 ERPs time-locked to stimulus onset from 13 scalp sites for pseudowords with early (solid) and late (dotted) deviation points. Note that stimulus onset is marked by the vertical calibration bar.

notice in Figs. 1 and 2 is that for both words and pseudowords with early recognition/deviation points (solid lines) the broad negative-going wave between 300 and 800 ms (what we will call the N400) begins and ends earlier in the ERP epoch than the comparable negativity for words and pseudowords with late recognition/deviation points (dotted lines). It appears as though this N400 is shifted approximately 100 ms to the right in late recognition/deviation point items.

To analyze this apparent shift in latency, the point of peak negativity between 300 and 750 ms was measured. ANOVAs on this measure revealed main effects of stimulus type (lateral: $F[1,29] = 24.62$, $p < 0.001$; midline: $F[1,29] = 5.71$, $p < 0.05$), and latency of recognition/deviation point (lateral: $F[1,29] = 73.26$, $p < 0.001$; midline: $F[1,29] = 37.15$, $p < 0.001$) suggesting that the negativity in the region of the N400 began sooner for words than for pseudowords and for stimuli with early than late recognition/deviation points. These effects were qualified by a significant stimulus type by latency to recognition/deviation point interaction (lateral: $F[1,29] = 10.63$, $p < 0.005$; midline: $F[1,29] = 4.62$, $p < 0.05$) indicating that the effect of latency to recognition/deviation point was larger for words than for pseudowords (compare Figs. 1 and 2). There was also an interaction between stimulus type and electrode site (lateral: $F[4,116] = 5.25$, $p < 0.005$) indicating that word/pseudoword differences tended to be larger at more anterior electrode sites.

2.2.2.2. Peak latency analysis from recognition point/deviation points. Grand average ERPs time-locked to the recognition/deviation points of words and pseudowords are plotted in Figs. 3 and 4. Note that in these figures the vertical calibration bar corresponds to the onset of recognition/deviation points rather than onset of the stimulus (as is Figs. 1 and 2).[4] The feature of these waveforms to focus on is the broad negativity occurring approximately at the calibration bars which roughly

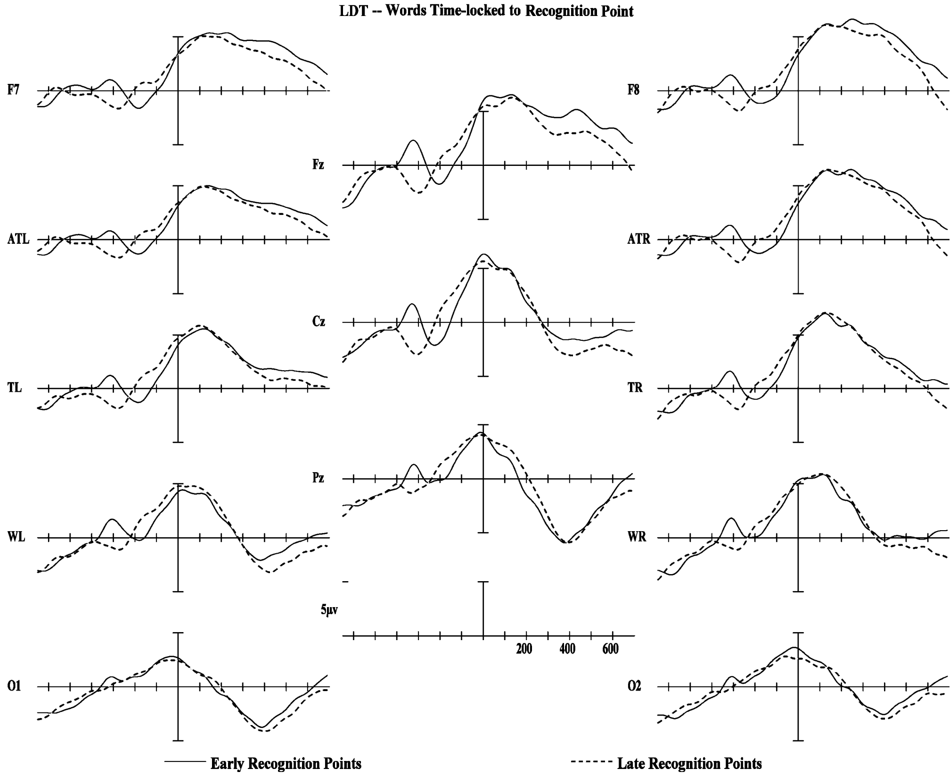

Fig. 3. Experiment 1 ERPs time-locked to uniqueness points in words with early (solid) and late (dotted) uniqueness points. Note that uniqueness point is marked by the vertical calibration bar in this figure.

[4] Note 650 ms of signal is displayed before each calibration bar. This region corresponds to the distance prior to the onset of recognition point/deviation points needed to reach a place in the EEG to obtain a baseline. We chose this point because the signal prior to the onset of recognition point/deviation points varied.

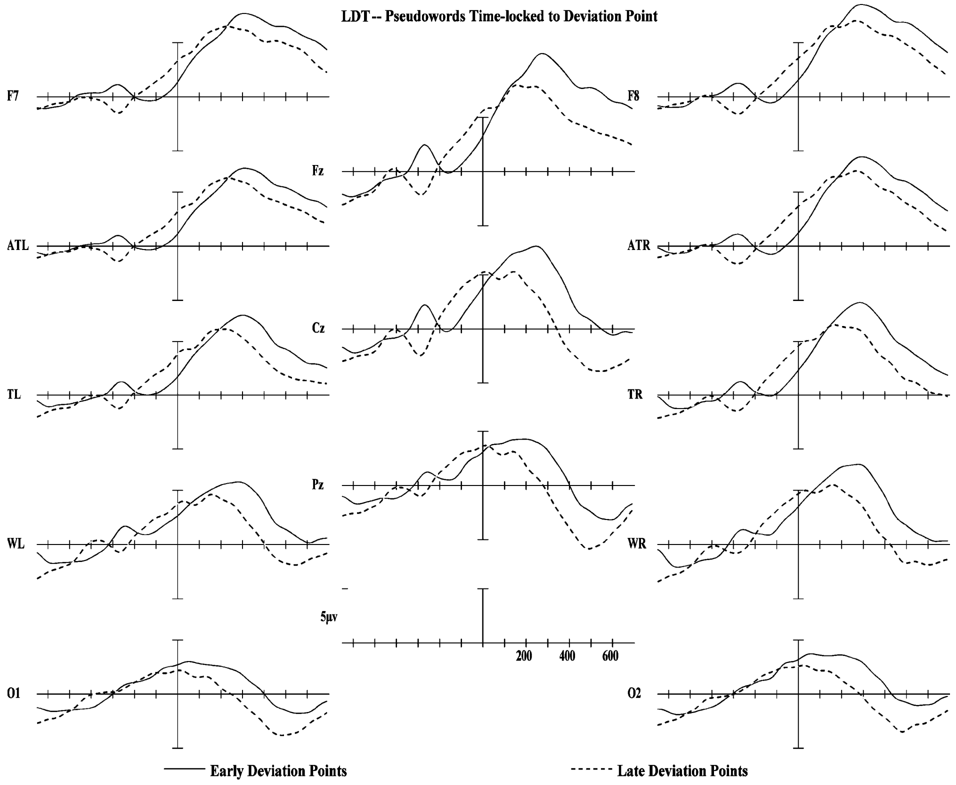

Fig. 4. Experiment 1 ERPs time-locked to deviation points in pseudowords with early (solid) and late (dotted) deviation points. Note that deviation point is marked by the vertical calibration bar in this figure.

corresponds to the late negativity seen later in onset-locked ERPs (cf. Figs. 1 and 2). Note that early and late recognition points for words seem to overlap in the region of the calibration bar in Fig. 3, while in Fig. 4 the negativity for early deviation points in pseudowords starts later than the negativity for late deviation points.

The peak latency analysis confirmed these observations. There were main effects for stimulus type (lateral: $F[1,29] = 31.28$, $p < 0.0001$; midline: $F[1,29] = 16.42$, $p < 0.0001$) and latency of recognition/deviation point (lateral: $F[1,29] = 24.26$, $p < 0.0001$; midline: $F[1,29] = 6.20$, $p < 0.05$) as well as a stimulus type by latency to recognition/deviation point interaction (lateral: $F[1,29] = 14.41$, $p < 0.001$; midline: $F[1,29] = 25.42$, $p < 0.0001$). Follow-up analyses revealed no difference in peak latency for words due to difference in recognition point (Fs < 1). However, pseudowords with early deviation points had significantly later N400s than those with late deviation points (lateral: $F[1,29] = 41.26$, $p < 0.0001$; midline: $F[1,29] = 25.88$, $p < 0.0001$).

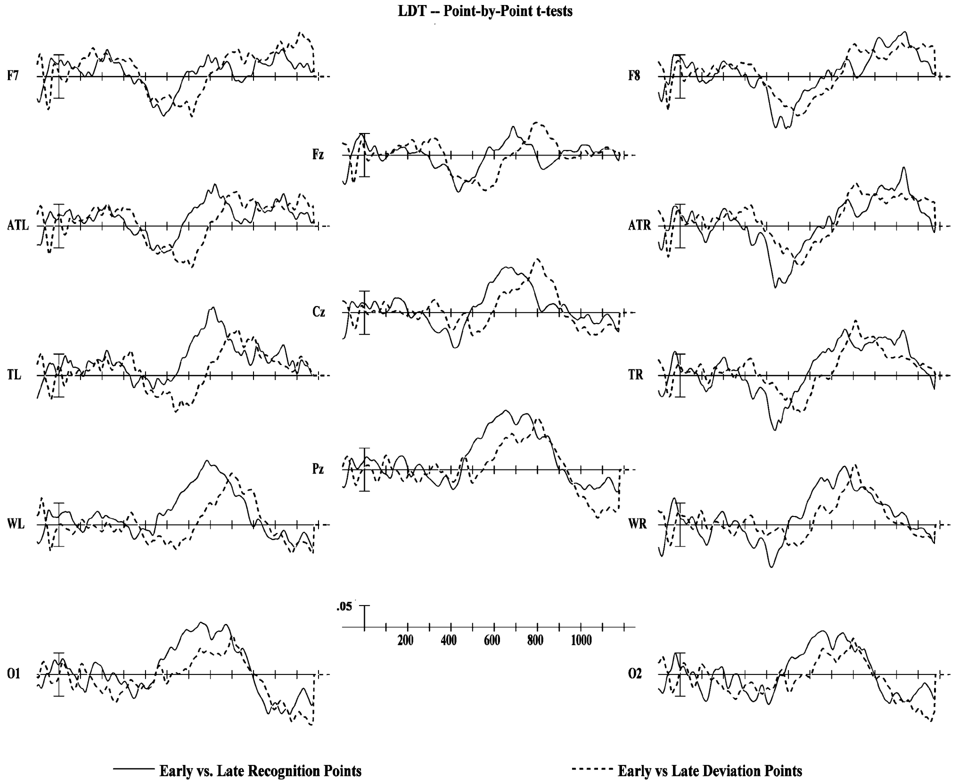

Fig. 5. Experiment 1 point-by-point t-tests contrasting early/late uniqueness point words and pseudowords time-locked to stimulus onset. Note that stimulus onset is marked by the vertical calibration bar which in this figure marks the 0.05 alpha level.

2.2.3. Time course analysis from stimulus onset

Peak latency analyses reveal trends in the global temporal characteristics of an ERP component. Because they focus on a single point in time they lack the precision necessary to quantify more fine grained temporal aspects of processing. The goal of these analyses was to evaluate the onset and offset of the N400. Fig. 5 is a graphical representation of the 256 consecutive t-tests at each electrode site contrasting early and late words (solid) and pseudowords (dotted). As with the ERP plots (e.g. Figs. 1 and 2) negative values (i.e. negative t values) are plotted in the upward direction. More positive t values reflect differences in which the early items are more negative than the late items. Negative t values reflect late more negative than early (note the calibration bar represents the alpha level of the t-test 0.05). Compare these graphs with Figs. 1 and 2 to gain a better understanding of regions of differences.

Fig. 5 shows a run (approximately 100 ms) of significant positive t values which began at about 400 ms at anterior sites for words and at about 500 ms for pseudowords. This suggests that early on in its time-course the N400 to early words and pseudowords was larger compared to late words and pseudowords. However, at about 550 ms for words and 650 ms for pseudowords, there was a long run of significant negative t values indicating that late items started producing significantly larger negativities than early items, especially at more posterior electrode sites. This pattern is consistent with the trend visible in the grand average ERPs plotted in Figs. 1 and 2, which suggest that the N400 started and ended sooner for early recognition/deviation point items. Another important aspect of this analysis is that the earliest reliable differences between early and late items occurred almost 200 ms before the offset of the shortest stimuli (400–450 ms words, 450–500 ms pseudowords).

2.2.4. Time course analysis from recognition point/deviation points

Fig. 6 shows the t-test plots for early and late words and pseudowords time-locked to recognition/deviation points. The primary goal of these analyses was to better quantify the differences apparent at the calibration bars in Figs. 3 and 4. Perhaps the most salient aspect of Fig. 6 is the large discrepancy in the number of significant tests for words and pseudowords starting from the recognition and deviation points. There was a lengthy period of significant positive t's for the early vs. late pseudoword contrasts starting as early as 50 ms after the deviation point and

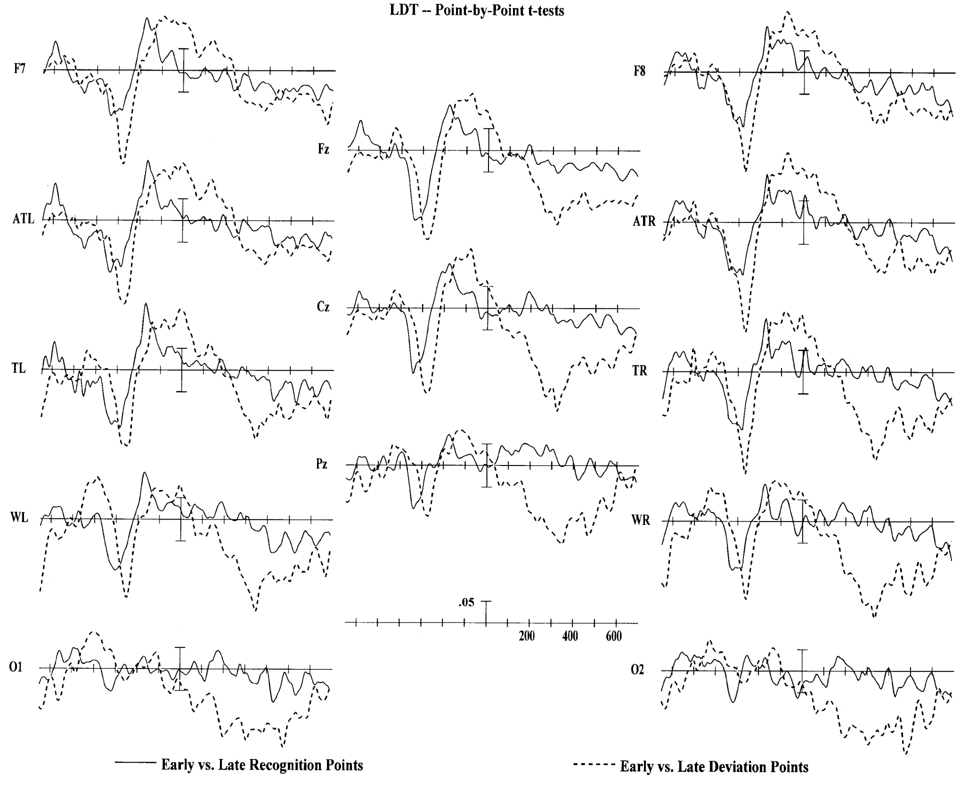

Fig. 6. Experiment 1 point-by-point t-tests contrasting early and late uniqueness point words (solid) and pseudowords (dotted) time-locked to uniqueness/deviation points.

lasting at some sites until 700 ms. This reflects the larger and later post-deviation point N400 for early than late deviation point pseudowords. However, no site showed more than three consecutive significant *t* values (positive or negative) after the recognition point for words.

This analysis also revealed significant differences in other windows. Differences between early and late words and pseudowords emerged in a string of significant negative *t* values that peaked at about 200 ms prior to the actual recognition points in words (i.e. to the left of the calibration bar—also see Fig. 3) and a bit later for pseudowords (about 100 ms pre-deviation point). These differences reflect the earlier rise of the N400 for late than early items, which is not surprising given that at this point about 100 ms more acoustic information has been heard for the late words. The large spike of significant positive *t* values peaking between 325 and 275 ms pre-recognition/deviation point primarily reflects differences in the timing of the exogenous N1 and P2 components for early and late items (see Figs. 3 and 4—due to recognition/deviation averaging the N1 and P2 for late items occurs about 100 ms sooner than for early items).

2.3. Discussion

N400 peak latency and RT measured from stimulus onset were both shorter for stimuli with recognition/deviation points occurring earlier in the acoustic signal than for stimuli with recognition/deviation points occurring later. ERP time-course analyses indicated that this displacement in peak latency was due to a more rapid rise and then fall of the N400 component for words and pseudowords with early than late recognition/deviation points. This pattern of finding is consistent with results from several previous studies (Marslen-Wilson, 1984; Goodman and Huttenlocher, 1988; Radeau et al., 1989; and Radeau and Morais, 1990) and fits with the overall predictions of theories of spoken word recognition such as the Cohort model which assert that early acoustic information plays a disproportionately important role during word recognition. This finding also supports earlier suggestions that N400 component may be a useful index of the temporal dynamics of spoken word processing (e.g. Holcomb and Neville, 1990; Soares et al., 1991; Woodward et al., 1990; Van Petten et al., 1999).

A second important finding was that RT and N400 latency time-locked to recognition points in words did not vary as a function of the location of recognition points being constant across early and late recognition points. Moreover, fine-grained time-course analyses of the ERPs indicated that while early and late words differed prior to their recognition point, after this point they did not differ. In other words consistent with the optimal efficiency prediction of the Cohort model, words with very different temporal properties and with very different initial ERP signatures, produced remarkably similar ERPs starting at the earliest possible point past which they could be differentiated from their lexical neighbors. These data extend the findings of Radeau et al. (1989) and Radeau and Morais (1990) who studied French words in gender classification and naming tasks to English words presented in a lexical decision task.

A third finding was that unlike the word results, RTs and N400 peak latencies time-locked to deviation points in pseudowords were significantly longer for items with early than late deviation points. Moreover, the ERP time-course analyses indicated that there were significant N400 onset and offset differences between early and late pseudowords with early pseudowords producing initially smaller and eventually larger N400s. Critically, the N400 offset lag for early pseudowords occurred well after the deviation point. This pattern is consistent with the findings of Goodman and Huttenlocher (1988) and Taft and Hambly's (1986) and supports their claim that listeners monitor acoustic information after recognition/deviation points during word processing. At the same time this finding is inconsistent with the claims of the Cohort model, which predicts that listeners should rapidly terminate lexical processing upon reaching a deviation point as there will no longer be any items left in the word initial cohort. Early and late pseudowords should receive comparable treatment past this point and therefore should produce similar neural and behavioral responses.

But what of the possibility that the lexical decision task encourages post-lexical strategies that might distort pseudoword responses in such a way that they do not accurately reflect the underlying processes of interest (i.e. word recognition)? We had proposed that if this were the case then there might be a dissociation between the RT and ERP results. One possible scenario was that RT would prove sensitive to strategic effects while the N400 would reveal the pure effects of word processing. A dissociation did not occur and both sets of dependent measures produced a similar pattern that is more consistent with the interpretation that post-deviation point information is monitored by listeners in the lexical decision task.

3. Experiment 2

Therefore, what can be concluded from the pattern of findings in experiment 1? On the one hand both the ERP and RT data from the word stimuli seem to clearly support the optimal efficiency predictions of Cohort model, while comparable data for the pseudowords seem to call this feature of the model into question. One possible explanation for this dissociation of word and pseudoword data is that the speeded lexical decision task artificially encouraged participants to prolong their processing of pseudowords thus resulting in differential early vs. late deviation point RTs and N400 time-courses. Conversely, the demands of lexical decision may have altered the normal word recognition process resulting in an unnatural focus on word recognition points. Normal spoken language processing may proceed very differently even when words are processed in isolation. For example, without the time pressure of making a speeded lexical decision listeners may delay their processing of spoken items until more information is available.

In experiment 2 participants were told to carefully pay attention to all spoken stimuli, but no explicit behavioral task was given. If the pattern of results obtained in Experiment 1 was due to the demands of the lexical decision task then the ERPs in this experiment should be quite different. In particular, if the Cohort model is correct, the time-course of the N400 to pseudowords should be comparable for early and late items. Conversely, without the processing demands of speeded lexical decision, the time-course of spoken word N400s might be delayed beyond recognition points resulting in a divergence between early and late items or a complete uncoupling of the recognition point/N400 time-course relationship. This would be evidence against the generalizability of the optimal efficiency prediction of the Cohort model.

Another motivation for experiment 2 was to reduce the influence of the P3 component on the ERPs elicited by words and pseudowords. The P3 can be a problem in studies focusing on the N400 because both components produce activity in the same temporal window thus resulting in component overlap. Moreover, because P3 peak latency can move earlier or later in the epoch depending on decision processes its can make unambiguous interpretation of the N400 results more difficult. By eliminating the binary lexical decision the influence of the decision P3 should also be reduced.

3.1. Methods

3.1.1. Participants

Thirty Tufts University undergraduates (13 female, 17 male) who were right handed (seven had left handed relatives in their immediate family) and native English speakers participated for course credit. All were between 18 and 22 years of age (mean = 19.9 years) and all had self-reported normal hearing.

3.1.2. Procedure

The testing procedure was the same as experiment 1 with the following changes. Participants were instructed to listen passively to each stimulus and were told that they simply needed to pay attention. Trials were initiated by participant' pressing a key. One second later a fixation point appeared in the center of a computer screen for 3.5 s. One second after the onset of the fixation point a stimulus was heard. Participants were asked not to blink while the fixation point was on the screen and to not press the button to advance the trial until after the fixation point disappeared from the screen. The minimum inter-trial interval was 3.5 s during which time participants saw a blank screen and were told they could blink. Participants were given a break approximately every 40 trials.

3.1.3. Data analysis

ERPs were averaged and analyzed as in experiment 1.

3.2. Results

3.2.1. Peak latency analysis from stimulus onset

Less than 10% of trials were rejected from the averaged ERP epochs for each condition because of artifact (mean number of trials in each average per participant; early words = 37, late words = 38, early pseudowords = 37, late pseudowords = 37).

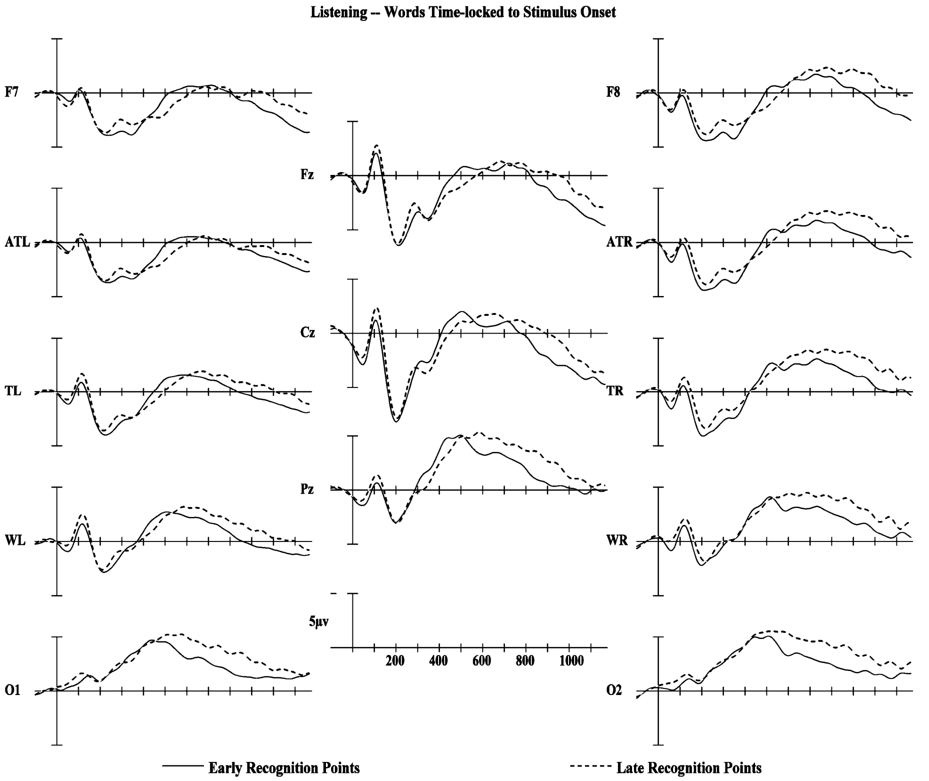

Fig. 7. Experiment 2 ERPs time-locked to stimulus onset from 13 scalp sites for words with early (solid) and late (dotted) uniqueness points. Note that stimulus onset is marked by the vertical calibration bar.

Fig. 7 displays grand average ERPs for words with early and late recognition points while Fig. 8 displays grand average ERPs for pseudowords with early and late deviation points. As in experiment 1, these waveforms include an initial central anterior maximum negativity with a peak latency near 100 ms (N1) followed by a positivity peaking near 200 ms (P2), followed by a broad negativity (N400). Unlike experiment 1 these waveforms do not include a large late positivity (P3). The relative absence of this component is likely due to participants not having had to make a binary classification and subsequent decision about the words in this experiment (Donchin and Coles, 1988). Note that Figs. 7 and 8 display the same shift towards longer latencies in the region of the N400 for stimuli with late recognition point/deviation points.

The peak latency analysis confirmed that the shift in the N400 towards longer latencies for stimuli with later recognition/deviation points was reliable (main effect stimulus type, lateral: $F[1,29] = 27.37$, $p < 0.001$; midline: $F[1,29] = 17.54$, $p < 0.05$; main effect latency to recognition point/deviation point, lateral: $F[1,29] = 31.47$, $p < 0.001$; midline: $F[1,29] = 17.28$, $p < 0.001$). Peak latency was earlier for words and stimuli with early recognition/deviation points.

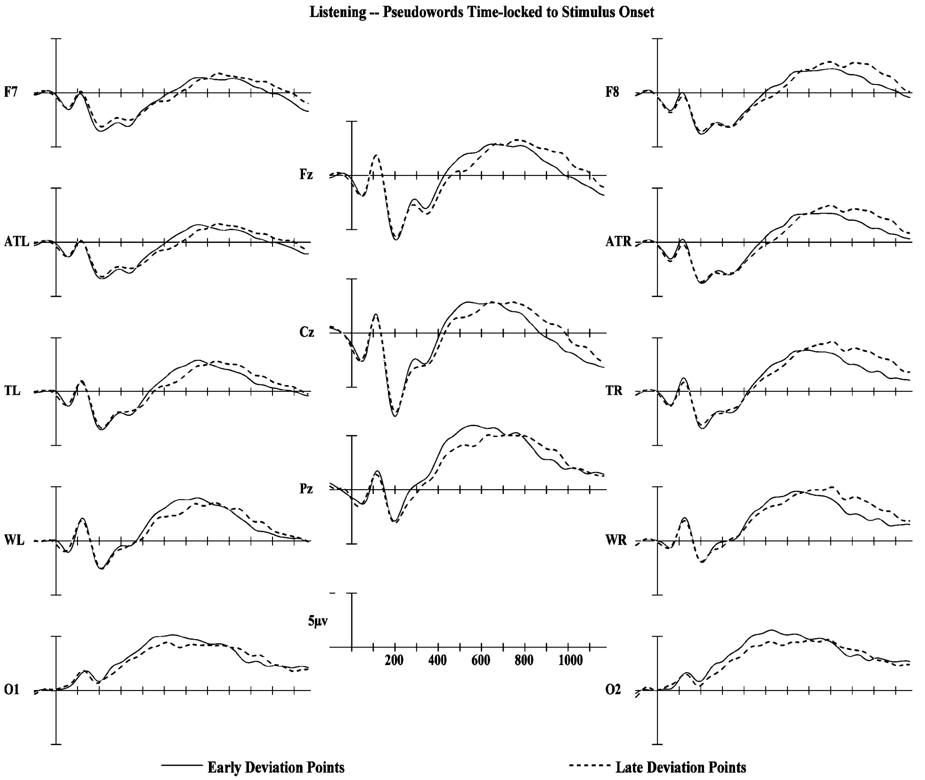

Fig. 8. Experiment 2 ERPs time-locked to stimulus onset from 13 scalp sites for pseudowords with early (solid) and late (dotted) deviation points. Note that stimulus onset is marked by the vertical calibration bar.

3.2.2. Peak latency analysis from recognition point/deviation points

Fig. 9 displays grand average ERPs time-locked to the recognition points of words. Fig. 10 displays grand average ERPs time-locked to the deviation points of pseudowords. Early and late recognition/deviation points are depicted with solid and dashed lines respectively. As in experiment 1, 650 ms prior to the onset of recognition point/deviation points were included in these waveforms (baselined −650 to −550). Focusing on the waveforms at the calibration bars, the same effects from Experiment 1 were apparent. The waveforms overlap when they are time-locked to recognition points and diverge when time-locked to deviation points. However, the latter effect is not nearly as dramatic as in experiment 1 and by the end of the epoch at a number of sites early and late pseudowords ERPs overlap.

Peak latency analyses provided support for these observations. The lateral analysis showed main effects for stimulus type (lateral: $F[1,29] = 52.60$, $p < 0.001$; midline: $F[1,29] = 44.99$, $p < 0.001$) and only marginal effect of latency to recognition/deviation point at lateral sites (lateral: $F[1,29] = 3.46$, $p = 0.073$). Peak latency was earlier for words and there was a trend for it to be later for stimuli with early

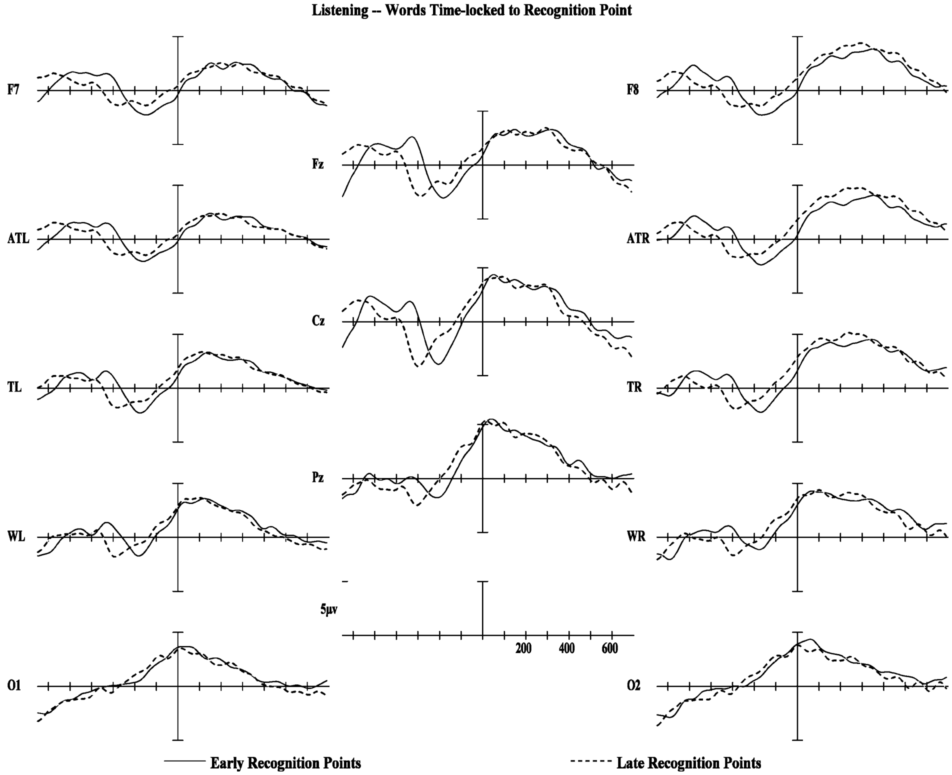

Fig. 9. Plotted in this figure are experiment 2 ERPs time locked to uniqueness points in words with early (solid) and late (dotted) uniqueness points. Note that uniqueness point is marked by the vertical calibration bar in this figure.

recognition/deviation points at lateral sites. Unlike experiment 1, there was not a significant stimulus type by latency to recognition/deviation point interaction ($F = 1.14$).

3.2.3. *Time course analysis from stimulus onset*

Fig. 11 shows the outcome at each electrode site within each of the 256 consecutive t-tests separately for words and for pseudowords. In the time band corresponding to the N400 (i.e. 300–750 ms) the majority of the differences were in the same direction as experiment 1 with runs of negative t values early (between 400 and 600 ms) and later runs of positive t values (between 700 and 1000 ms).

3.2.4. *Time course analysis from recognition point/deviation points*

Fig. 12 shows the consecutive early versus late t-test results for words and pseudowords time-locked to the onset of recognition point/deviation points. The goal here was to better quantify the differences apparent surrounding the calibration bars in Figs. 9 and 10. As in experiment 1, while there was a run of 100 ms or so of

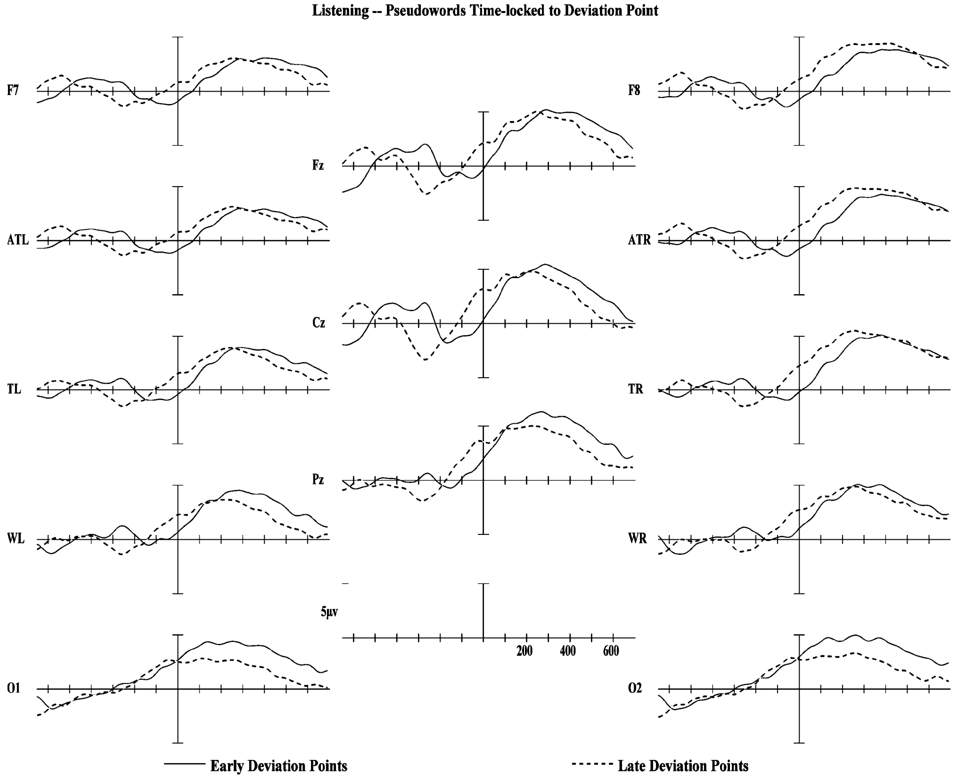

Fig. 10. Plotted in this figure are experiment 2 ERPs time locked to deviation points in pseudowords with early (solid) and late (dotted) deviation points. Note that deviation point is marked by the vertical calibration bar in this figure.

significant negative *t* values at most sites prior to the recognition points in words (indicating that late words were significantly more negative going than early words), very few of the comparisons between early and late words were significant after the recognition point. In fact, no site produced more than three consecutive significant responses after the recognition point. Pseudowords, on the other hand, produced a large number of significant negative *t* values both before and after their deviation points indicating that the initially larger negativity to late items did not subside until well after the deviation point. However, unlike experiment 1, there was less evidence of a later run of significant positive *t* values, which would indicate that the early items eventually overtook the late items and ended up with larger N400s. Only at the posterior sites was there evidence of a run of positive t values.

3.3. Discussion

Experiment 2 was a replication of experiment 1 using the same stimulus materials but a different task. Rather than making explicit lexical decisions participants were

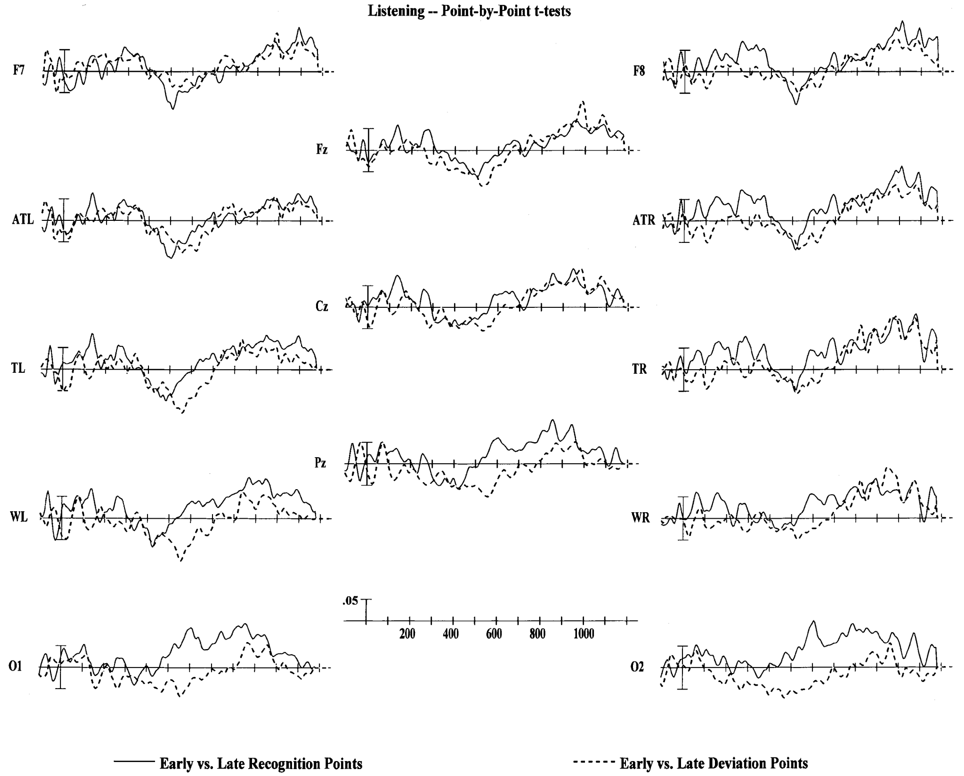

Fig. 11. Experiment 2 point-by-point t-tests contrasting early and late uniqueness points words (solid) and pseudowords (dotted) time-locked to stimulus onset.

asked to attentively listen to all spoken stimuli, but were told to make no overt behavioral response. As a result, only ERP measures were collected.

The findings of this experiment replicate those of the first experiment in many of the most important ways. First, when ERP recordings were time-locked to word/pseudoword onset the latency of the N400 and its time-course were earlier for words and pseudowords with early recognition/deviation points than similar items with later recognition/deviation points. Second, when time-locked to recognition points N400 latency and N400 time-course were equivalent for early and late words. These data support the contention that the results obtained with words in experiment 1 were not due to the demands of making speeded explicit lexical decisions and reaffirms in a more natural word processing paradigm that spoken words are processed in a manner consistent with the predictions of the Cohort model.

Early and late pseudowords produced mixed results with regards to the pattern of effects seen in experiment 1. In the analysis of N400 peak latency there was no statistically reliable evidence that early and late pseudowords differed when ERPs were time-locked to deviation points. This is unlike experiment 1, where late pseudowords produced significantly earlier N400s than early pseudowords. How-

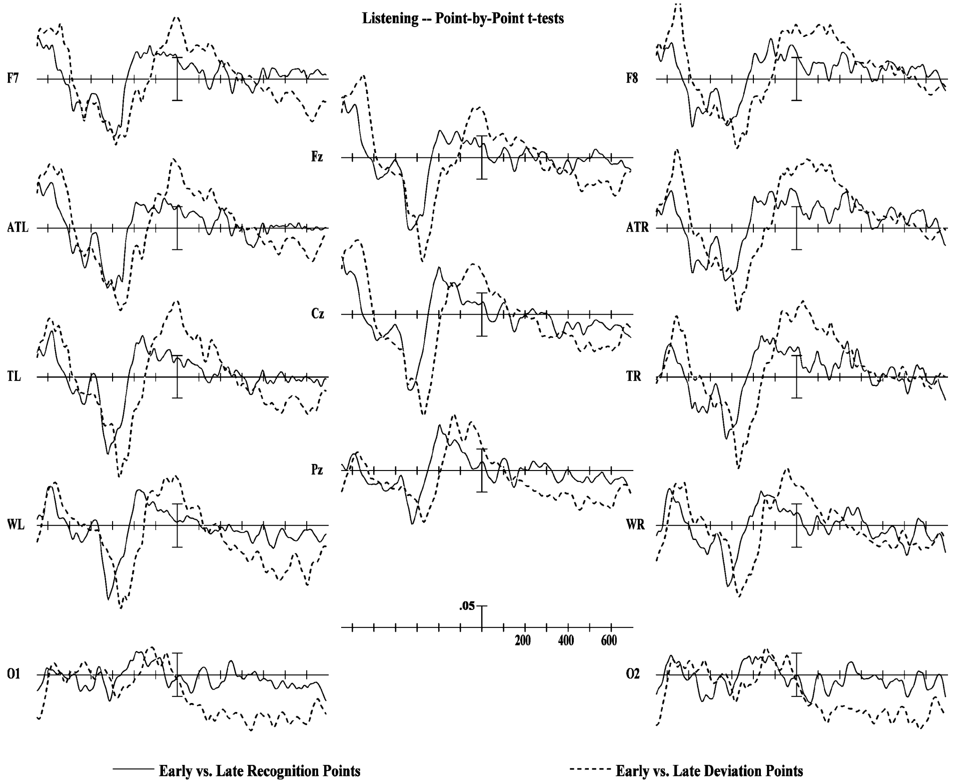

Fig. 12. Experiment 2 point-by-point t-tests contrasting early and late uniqueness point words (solid) and pseudowords (dotted) time-locked to uniqueness/deviation points.

ever, the deviation point time-locked time-course analysis did reveal that early pseudowords produced a slower rising N400 response than late pseudowords, and unlike words this effect spilled over into the post-deviation point epoch. However, the large difference in N400 offset latency seen in Experiment 1 for early and late pseudowords was not as apparent in experiment 2. Only the most posterior sites produced evidence of a significantly later offset of the N400 for early pseudowords. Therefore, the evidence supporting post-deviation point processing is relatively weaker in this experiment as are the conclusions that can be drawn for the importance of this effect for normal word processing.

One possibility for this weak support is that without the requirement of an explicit lexical decision participants simply paid less attention to information following the deviation point. However, that participants paid attention to the stimuli in general is clear given the similarity of word effects to those seen in experiment 1.

4. General discussion

The primary goal of this study was to test the optimal efficiency prediction of the Cohort model. The optimal efficiency prediction states that spoken words are frequently recognized (selected) at the first point past which no other word is compatible with the incoming acoustic signal. While there has been some previous support for this proposal, especially with respect to real word processing (e.g. Radeau et al., 1989), at least two prior studies have suggested that certain types of pseudowords may be processed past the point where they are clearly no longer potential words (Taft and Hambly, 1986; Goodman and Huttenlocher, 1988). This is problematic for the Cohort model because Marslen-Wilson himself used pseudo-word data from the lexical decision task to argue for optimally efficient word processing (see Marslen-Wilson, 1984).

In the current study we systematically manipulated the time point in the acoustic signal past which an item was uniquely specified as a word (i.e. the recognition point) or was no longer a possible word (i.e. the deviation point). The peak latency and time-course of the N400 component (experiments 1 and 2) and RT (experiment 1) were measured both from word/pseudoword onset and from the onset of recognition and deviation points.

In experiment 1 RT and the time-course of the N400 were earlier for words and pseudowords with early than late recognition/deviation points. However, when time-locked to recognition/deviation points there was a dissociation of the word and pseudoword effects. While N400 latency (peak and time-course) and RT were constant for early and late recognition pointwords, pseudowords with early deviation points produced ERPs with later N400s and slower RTs than pseudowords with late deviation points. This result is consistent with the pattern reported by Goodman and Huttenlocher (1988) from which they argued against the viability of the optimal efficiency hypothesis.

Experiment 2 used a passive listening task where participants were told to attend to the spoken stimuli, but were not required to make overt lexical decisions. This experiment sought to determine if the lexical decision task itself was responsible for the pattern of results in experiment 1. Particularly, whether making speeded lexical decisions was somehow responsible for the similar N400s for early and late real words or for the later N400s for early than late pseudowords. The word ERP data from experiment 2 closely paralleled those from experiment 1. Most importantly early and late words had very similar N400s when time-locked to recognition points. However, the early/late pseudoword N400 dissociation found in Experiment 1 was not reliable in experiment 2.

Several points can be made about these findings. First, the data from the real words in both experiments strongly support the optimal efficiency prediction of the Cohort model. Even in a passively processed word list, ERPs time-locked to recognition points demonstrated a remarkably constant N400 latency and time-course for early and late words. Moreover, this effect was accompanied in Experiment 1 by a similar pattern for RT. Participants were equivalently fast at making their word decisions for early and late words when responses were time-locked to recognition points.

Second, the ERP data add another dimension of support to the Cohort model, in particular, to the validity of the concept of recognition points. In both experiments, prior to the recognition point, late words produced a faster raising N400 than the early words. This is not surprising given that the late words onset, on average, 100 ms sooner than the early words. What is noteworthy is that this faster raise time did not translate into a faster decay time—the N400 was simply not displaced 100 ms earlier for late words. Rather, after the recognition point, both early and late words produced N400s with a similar decay function. This can be seen most clearly in the time-course analyses (Figs. 6 and 12) where there was a relatively long run of significant differences between early and late words prior to and right up to the recognition point, but there were virtually no significant differences after this point. If participants were differentially processing words after the recognition point, which is the claim of Taft and Hambly (1986) and Goodman and Huttenlocher (1988), it is not obvious from the lexical decision RTs in Experiment 1 or from the ERPs in either experiment, which were recorded for over 700 ms after the recognition points.

Support for the validity of recognition points should not be confused with stronger claims such as word processing ending with recognition points. Neither our RT or ERP data support this position. This is because RT occurred almost 500 ms after the recognition points, which would seem to be ample time for additional word processing. And while the N400s for early and late words did not differ after the recognition point, the N400s to both types of items had not run its course for as long as another 500 ms (see Figs. 4 and 9).

Third, the pseudoword results were not as clear cut. There was strong RT and ERP evidence from experiment 1 that pseudowords continue to be differentially processed after their deviation points. However, experiment 2 did not fully support this result and found a weaker trend for post-deviation point differences. One possibility for this different pattern of results between experiments is that the pseudoword effects in experiment 1 may have been due entirely, or in part, to the demands of the lexical decision task. The presence of a similar, but weaker trend in experiment 2 could have been the result of participants making 'covert' lexical decisions on a subset of trials, even though the task was passive listening (the presence of pseudowords in the lists could have encouraged such a strategy). In other words, the experiment 2 pseudoword pattern may have resulted from a mixture of trials; some where participants made covert lexical decisions which resulted in a delayed N400s for early deviation point items (the experiment 1 pattern) and others where no lexical decision was made and where there was no difference in the time course of the N400 for early and late pseudowords (i.e. the pattern seen for words).

But why should pseudowords be processed so differently from words? One explanation is that while word recognition may be typically associated with recognition points, this process may nevertheless require some additional time, perhaps as much as 500 ms to be fully completed. (Note that this would not violate optimal efficiency if processing is pipelined; i.e. if the next word can enter the pipeline once the current word has reached the 'recognition' point.) According to this view in many, perhaps most cases, word recognition processes would run to completion unheeded from this point. However, the presence of items with ambiguous information past this point might alter or disrupt normal post-recognition point processing. And it might be this disruption of the normal process that differentially affected the time-course of the N400 (and RT) for early and late pseudowords.

But why should this happen if normal word recognition can proceed from an early point just as efficiently as from a later point? Should not participant, as suggested by Marslen-Wilson's (1984) results, have been able to use the non-lexical information occurring after this point in pseudowords to classify these items as nonwords? In the case of relatively radical disruptions this might be possible. But if disruption is less severe, as in the current study and in Goodman and Huttenlocher (1988) then maybe the disruption signals the need for the listener to use a different process. In the real world such minor disruptions indicate that the listener has not clearly heard a word (e.g. it was mispronounced or a phoneme was deleted due to ambient noise). It is apparent that listeners can readily deal with this type of 'noise'. But Marslen-Wilson has made it clear that the Cohort model was not meant to account for what happens when the system encounters such noise. In other words, the Cohort model is a theory of first pass processing and in that sense has little to say about the processes that are engaged when the normal word recognition process fails. Presumably under noise conditions some other top-down process aids in recovering from disruptions.

Taft and Hambly (1986) proposed such a post-lexical, top-down mechanism to account for their pseudoword data. According to this account ambiguous items cannot be fully processed until all acoustic information is available. At this point they are checked against lexical items activated via bottom-up processing. Such a mechanism would seem to account for our pseudoword data in experiment 1 as well. Since late pseudowords ended sooner than early pseudowords (when time-locked to deviation points) these would have on average been checked sooner resulting in faster RTs and earlier N400s. However, Taft and Hambly implied that this mechanism also works during normal word recognition. There are two problems with this conclusion. First, our word RT and N400 time-course data strongly suggest that listeners do not differentially process early and late words after their recognition points. Second, such a system would probably be too slow and demand too many central resources to subserve normal 'fast' conversational word recognition.

In summary, the pattern of findings from this study with words is most consistent with the view that lexical decision RT and N400 time-course initially reflect normal fast spoken word processing. However, the pseudoword data suggest that these measures are also sensitive to a later top-down process. Evidence for this possibility can be seen in the comparison of the word and pseudoword ERPs from this study. While similar during the rising phase of the N400, the ERPs to pseudowords ended up quite different from those to words starting at the N400 peak and continuing into the declining phase of this component. Moreover, there is evidence that the late word/pseudoword difference may emanate from a different neural source than that producing the overall shift in N400 time-course. This can be seen on the left side of Fig. 13, which contrasts all words and pseudowords from Experiment 1. In addition

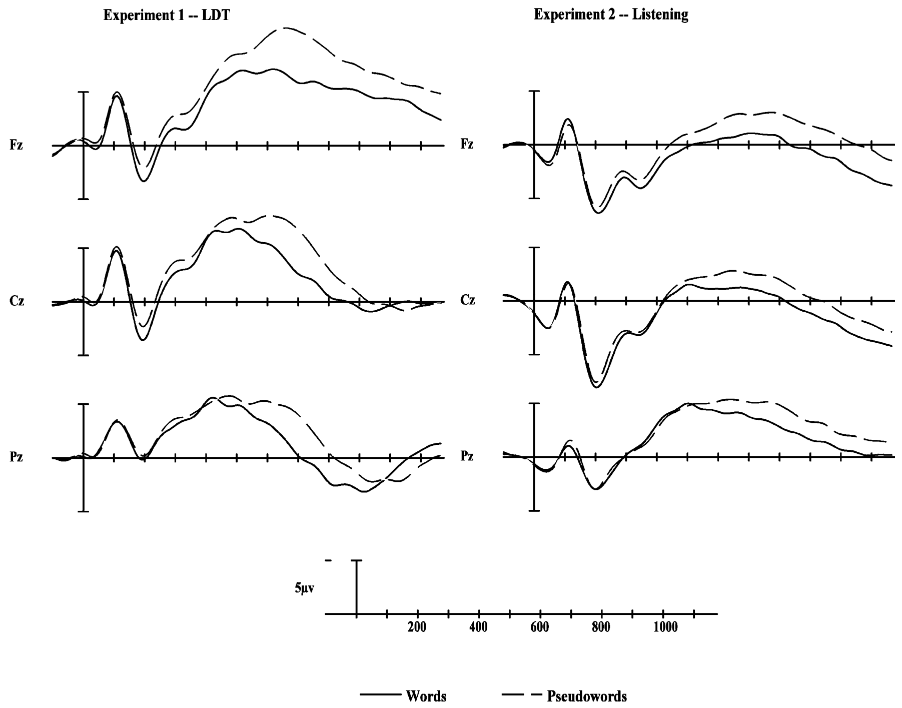

Fig. 13. ERPs time-locked to stimulus onset from three midline scalp sites for all words (solid) and all pseudowords (dotted) from experiments 1 (left) and 2 (right). Note that stimulus onset is marked by the vertical calibration bar.

to the generally later offset of the N400 for pseudowords compared to words (seen at all electrode sites), at the more anterior sites pseudowords produced a larger late negativity than words (e.g. compare Fz and Pz between 600 and 800 ms). This pattern has been reported in previous lexical decision studies (e.g. Holcomb and Neville, 1990) and suggests that pseudowords undergo additional processing, possibly in a separate system. This interpretation is consistent with the view that the N400 is not a monolithic component, but rather is composed of several distinct sub-components (e.g. Nobre and McCarthy, 1994; Kounios and Holcomb, 1994). In the current context this suggests that the relatively early, broadly distributed N400 activity reflects rapid bottom-up word recognition processes, while the later more anterior negativity reflects a subsequent top-down checking process. Consistent with this view the analogous data to those shown on the left side of Fig. 13 for experiment 1 also show that pseudowords elicited a larger N400 then words in experiment 2 (Fig. 13, right side). However, importantly, the absolute size of the difference was not as large and the anterior/posterior differences were not so apparent. One possibility for the reduction in the word/pseudoword differences for experiment 2, especially at anterior sites may reflect participants having made no overt and relatively fewer covert lexical decisions in this experiment and therefore having engaged post-lexical checking less frequently.

Confirmation of this account of word/pseudoword differences in the N400 time window will have to await future studies that more carefully control the degree participants need to engage such a process. For example, it may be that certain words (e.g. those of extremely low frequency or familiarity) may also require this type of check, in which case they too should elicit the more anterior late negativity.

Acknowledgements

This research was supported by HD25889 to P.J.H.

References

Bentin, S., McCarthy, G., Wood, C.C., 1985. Event-related potentials, lexical decision and semantic priming. Electroencephalogr. Clin. Neuropsychol. 60, 343–355.

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

Brown, C., Hagoort, P., 1993. The processing nature of the N400: evidence from masked priming. J. Cognit. Neurosci. 5, 34–44.

Chumbley, J.I., Balota, D.A., 1984. A word's meaning affects the decision in lexical decision. Mem. Cogn. 12, 629–649.

Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating? Behav. Brain Sci. 11, 355–372.

Francis, W.N., Kucera, H., 1982. Frequency Analysis of English Usage. Houghton Mifflin Co, Boston, MA.

Goodman, J., Huttenlocher, J., 1988. Do we know how people identify spoken words? J. Mem. Lang. 27, 684–698.

Grosjean, F., 1980. Spoken work recognition processes and the gating paradigm. Percept. Psychophys. 28, 267–283.

Grosjean, F., 1985. The recognition of words after their acoustic offset: Evidence and implications. Percept. Psychophys. 38, 299–310.

Guthrie, D., Buchwald, J.S., 1991. Significance testing in differences potentials. Psychophysiology 28, 240–244.

Holcomb, P.J., 1988. Automatic and attentional processing: An event-related brain potential analysis of semantic priming. Brain and Language 35, 66–85.

Holcomb, P.J., 1993. Semantic priming and stimulus degradation: implications for the role of the N400 in language processing. Psychophysiology 30, 47–61.

Holcomb, P.J., Neville, H.J., 1990. Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. Lang. Cognit. Process. 5, 281–312.

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

Holcomb, P.J., O'Rourke, T., Grainger, J. (in press). An event-related brain potential study of orthographic similarity. J. Cognit. Neurosci.

Kounios, J., Holcomb, P.J., 1992. Structure and process in semantic memory: evidence from event-related brain potentials and reaction time. J. Exp. Psychol. Gen. 121, 459–479.

Kounios, J., Holcomb, P.J., 1994. Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. J. Exp. Psychol. Learn. Mem. Cogn. 20, 804–823.

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

Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307, 161–163.
Marslen-Wilson, W.D., 1984. Function and process in spoken word recognition. In: Bouma, H., Bouwhuis, D. (Eds.), Attention and Performance: Control of Language Processes. Erlbaum, Hillsdale, NJ, pp. 125–150.
Marslen-Wilson, W.D., 1987. Functional parallelism in spoken word recognition. In: Frauenfelder, U.H., Tyler, L.K. (Eds.), Spoken Word Recognition. The MIT Press, Cambridge, MA, pp. 71–103.
Marslen-Wilson, W.D., 1990. Activation, competition and frequency in lexical access. In: Altmann, G.T.M. (Ed.), Cognitive Models of Speech Processing: Psycholinguistic and Computational Perspectives. MIT Press, Cambridge, Ma, pp. 148–172.
Marslen-Wilson, W.D., 1993. Issues of process and representation in lexical access. In: Atlmann, G.T.M., Shillrock, R. (Eds.), Cognitive Models of Speech Processing: The Second Sperlonga Meeting. Lawrence Erlbaum, Hove, UK, pp. 187–210.
Marslen-Wilson, W.G., Tyler, L.K. 1980. The temporal structure of spoken language understanding. Cognition, 8, 1–71.
Marslen-Wilson, M.D., Welsh, A., 1978. Processing interactions and lexical access during word recognition and continuous speech. Cognit. Psychol. 10, 29–64.
Meyer, D., Schvaneveldt, R. 1971. Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. J. Exp. Psychol., 90, 227–234.
Nobre, A.C., McCarthy, G., 1994. Language-related ERPs: scalp distributions and modulation by word type and semantic priming. J. Cognit. Neurosci. 6, 233–255.
Nusbaum, H.C., Pisoni, D.B., Davis, C.K., 1984. Sizing uniqueness point the hoosier mental lexicon: Measuring the familiarity of 20,000 words. In: Research on Speech Perception Progress Rep. No. 10. Speech Research Laboratory, Department of Psychology, Indiana University, Bloomington, IN.
Radeau, M., Morais, J., 1990. The uniqueness point effect in the shadowing of spoken words. Speech Commun. 9, 155–164.
Radeau, M., Mousty, P., Bertelson, P., 1989. The effect of the uniqueness point in spoken-word recognition. Psychol. Res. 51, 123–138.
Rugg, M.D., 1987. Dissociation of semantic priming, word and non-word repetition effects by event-related potentials. Q. J. Exp. Psychol. 39A, 123–148.
Soares, I., Collet, L., Duclaux, R., 1991. Electrophysiological correlates of auditory lexical decision: an attempt to test the 'Cohort model'. Int. J. Neurosci. 57, 111–122.
Taft, M., Hambly, G., 1986. Exploring the Cohort Model of spoken word recognition. Cognition 22, 259–282.
Tyler, L.K., Frauenfelder, U.H., 1987. The process of spoken word recognition: an introduction. Cognition 25, 1–20.
Van Petten, C., Coulson, S., Rubin, S., Plante, E., Parks, M., 1999. Time course of word identification and semantic integration in spoken language. J. Exp. Psychol. Learn. Mem. Cogn. 25, 394–417.
Woodward, S.H., Owens, J., Thompson, L.W., 1990. Word-to-word variation in ERP component latencies: spoken words. Brain Lang. 38, 488–503.