Published in final edited form as:

Neuroreport. 2018 November 07; 29(16): 1379–1383. doi:10.1097/WNR.0000000000001121.

The use of acoustic information in lexical ambiguity resolution: An ERP study

Stephanie C. Leach and Erin Conwell

Department of Psychology, North Dakota State University

Abstract

Words that can be used as both noun and verb create regions of syntactic ambiguity that could create processing challenges for listeners. However, acoustic properties, such as duration, differ between noun and verb uses of such words [1–4], and listeners may use these differences to facilitate ambiguity processing. In this study, we replaced noun uses of ambiguous words with verb uses to determine whether these manipulations affected the N400 event related potential (ERP), which is associated with semantic violations, or the P600 component, which is associated with syntactic ambiguity. The results suggest that the acoustic differences between noun/verb polysemes mitigate the extent to which these words are perceived as ambiguous, although the results do not indicate whether replacing one with the other produces a meaning violation. Durational differences in noun/verb polysemes may affect their processing in fluent speech.

Keywords

Lexical ambiguity; noun/verb homophones; ERP

Introduction

Despite widespread ambiguity in language, listeners show few difficulties in interpreting speech. For example, words that are both nouns and verbs create regions of lexical and syntactic ambiguity, but do not tend to create confusion in listeners. One possible reason that these words do not create processing problems may be that noun uses are consistently longer than verb uses of such *ambicategorical* words [1–4]. Therefore, listeners could use duration to make decisions about words in ambiguous contexts, even before they have access to higher-level cues such as sentence meaning or structure. In previous behavioral research, participants used phonological cues to determine the grammatical category of written words [5]. Furthermore, isolated auditory exemplars of ambicategorical words produced differential neural activity [6–9]. However, no research has yet examined whether the acoustic differences between noun and verb uses of ambiguous words are used by listeners during online speech processing. The consistent presence of acoustic information does not necessarily indicate that the information is used by listeners to resolve ambiguities.

This study used event related potentials (ERPs) to ask whether durational cues affect processing of ambiguous words in fluent speech. The N400, which indicates a semantic violation [10], is a broad component that typically peaks around 400 ms post-stimulus onset. The P600, which occurs regions of ambiguity in sentences, including ambiguous words [11], is a positive-going component that peaks approximately 600 ms post-stimulus onset. If a verb use placed in a noun context (or vice versa) creates a semantic violation, we should see an enhanced N400 in response to such a replacement. Because ambiguous words produce an enhanced P600, we should see a strong P600 to the target words. If, however, participants perceive auditory exemplars of ambicategorical words as unambiguous because of their durations, then we should not see a P600 to these words. Furthermore, although the P600 is also reported relative to regions of syntactic ambiguity, the sentences in this study are only syntactically ambiguous because of the grammatical category ambiguity of the target words. Therefore, if the target word is disambiguated by its acoustics, the syntactic ambiguity disappears.

Method

Participants

Participants were 18 right-handed native English speakers (10 female; mean age 18.9 years, standard deviation 1.47) who received course credit for participation. Eight additional participants were excluded due to recording problems (6) or handedness (2). Participants gave informed consent prior to beginning the study.

Stimuli

The stimuli included some items used in a previous study [2] as well as additional stimuli that were constructed for this study. Twelve target words (*book, cook, look, push, dig, kick, kiss, tip, chat, nap, pack, snack*) were each used in 6 pairs of noun and verb contexts. Context pairs were acoustically identical for at least two syllables following the target word. A female, native English speaker recorded the sentences which were then edited as follows: each target word was replaced once with a use of that word produced in the same grammatical category (condition: *same*) and once with a use of that word produced in the other category (condition: *different*). A paired t-test showed that noun uses were longer than verb uses ($M_{\text{noun}}=0.328$; $M_{\text{verb}}=0.309$; $t(71)=3.22$, $p<0.001$). The final stimulus set contained 288 target sentences and 144 additional sentences containing non-words. A complete list of target stimuli is available online at <https://osf.io/tj2gh/>.

Design

Stimuli were presented in four blocks. *Noun same* blocks consisted of sentences where a noun usage of a target word was spliced into a noun context. *Noun different* blocks consisted of sentences where a verb usage of a target word was spliced into a noun context. *Verb same* blocks consisted of sentences where a verb usage of a target word was spliced into a verb context. *Verb different* blocks consisted of sentences where a noun usage of a target word was spliced into a verb context. Each block contained 72 target sentences and 18 sentences with non-words, which were presented in random order. Blocks were each presented twice,

for a total of eight blocks in the study. Presentation of blocks was counter-balanced across participants. Participants heard a total of 720 trials in the study.

Procedure

Neural activity was recorded with a 64-channel EGI HydroCel Sensor Net v2.0 (Electrical Geodesics, Inc., Eugene, OR). Once the electrode net was applied to a participant's head, they were seated about 50 cm away from a pair of speakers in a dimly lit, electrically shielded, soundproof room. Participants were asked to refrain from blinking or moving their eyes while audio was playing. To ensure that participants attended to the stimuli, they were instructed to hit a button on a response box when they heard a non-word and then wait for the next trial to begin. If all the words in the production were real words, participants were to simply wait for the next trial to begin. The interstimulus interval was 1000 ms.

ERP recording and analysis

Neural activity was continuously recorded during the study. The Cz electrode was the reference electrode. The EEG signal was amplified by an EGI NetAmps 200 amplifier with a sampling rate of 250 Hz and a bandpass filter of 0.1–100 Hz. All impedance levels were below 100 k Ω at the start of the experiment. All data collection was performed in compliance with a protocol approved by the NDSU Institutional Review Board.

The EEG data were processed with NetStation 4.3.1 (EGI, Eugene, OR) and lowpass filtered at 30 Hz. Two 900 ms analysis windows were created. The first began at target word offset and the other started at the offset of the word that disambiguated the category of the target word. Each 900 ms segment was baselined using the 100 ms segment immediately preceding it. The average voltage in the baseline period was subtracted from the average voltage in the post-offset period. After this baseline adjustment, a custom program detected and removed movement artifacts. The data for each participant were averaged within each condition: *noun same*, *noun different*, *verb same*, *verb different*, and *non-word*. Trials containing artifacts were excluded. Condition averages were then re-referenced to an average reference.

Recordings from 4 electrodes over front-central sites (Fz, AFz, F1, and F2) were averaged for analysis. These electrodes were selected because previous research indicates that this region produces the N400 when a semantic violation is present [12] and because an N400 to non-words could be clearly identified in these sensors. Visual inspection of the data revealed a clear but delayed N400 to non-words peaking at 580 ms following non-word offset (Fig. 1). Therefore, we adjusted the time windows for analysis to reflect this delay. The time window for the N400 analyses was 480–680 ms following stimulus offset and the time window for the P600 analyses was 680–880 ms following stimulus offset. All reported comparisons used the peak amplitude from these time windows after the baseline adjustment.

The data were analyzed using linear mixed effects models implemented in R [13] using the lme4 package [14]. Type of replacement (same or different category), lexical category indicated by the syntax (noun or verb), and the interaction of the two were included as fixed factors, while participant was the random effect. The dependent variable was peak amplitude

in the analysis window. Statistical significance was based on degrees of freedom estimated using the Satterthwaite approximation using the lmerTest extension [15].

Results

Target Word Analyses

The mean voltage in each condition over the 900 ms window following target word offset is presented in Fig. 1. Analysis of the N400 component was conducted using peak amplitude by participant in the time window 480–680 ms post-target word offset. The analyses showed no main effect of grammatical class ($t(54)=1.84$, $p=0.07$), but a significant main effect of replacement type ($t(54)=2.09$, $p=0.04$), which reflects greater negativity to the different replacement condition than to the same replacement condition. This analysis also showed a significant interaction of replacement type and grammatical class ($t(54)=-2.59$, $p=0.01$), due to increased negativity to words that had been produced as verbs relative to those produced as nouns. Complete results of the mixed model are in Table 1.

Analysis of the P600 component was conducted using peak amplitude by participant in the time window 680–880 ms following target word offset. This analysis showed no main effect of either grammatical class ($t(54)=1.36$, $p=0.18$) or replacement type ($t(54)=1.22$, $p=0.25$). There was a significant interaction between the two factors ($t(54)=-2.09$, $p=0.04$), which resulted from greater negativity to words produced in verb contexts than to those produced as nouns. Complete results of the mixed model are in Table 2.

Disambiguation Point Analyses

The mean voltage in each condition over the 900 ms window following disambiguating word offset is presented in Fig. 2. Analysis of the N400 component was conducted using peak amplitude by participant in the time window 480–680 ms post-disambiguation. The analyses showed no main effect of grammatical class ($t(54)=1.36$, $p=0.18$), nor of replacement type ($t(54)=1.22$, $p=0.23$), and no interaction of replacement type and grammatical class ($t(54)=-1.42$, $p=0.16$). Complete results of the mixed model are in Table 3. Analysis of the P600 component was conducted using peak amplitude by participant in the time window 680–880 ms following target word offset. This analysis also showed no main effect of grammatical class ($t(54)=1.47$, $p=0.15$), no main effect of replacement type ($t(54)=1.22$, $p=0.23$), and no interaction between the two factors ($t(54)=-1.22$, $p=0.23$). Complete results of the mixed model are in Table 4.

Discussion

The goal of this study was to determine whether listeners use the phonetic differences between noun and verb uses of polysemes to resolve ambiguity in fluent speech. We found that replacing a noun use with a verb use (and vice versa) resulted in a more negative EEG response in the N400 window than replacing a noun use with a noun use (or a verb use with a verb use) did. This indicated that using an exemplar that had the “wrong” acoustic properties may have created a semantic violation, even though the two exemplars were phonemically identical. We found no P600 response at the target word or disambiguation point as a function of condition. Because the P600 is associated with syntactic ambiguity, its

absence at the disambiguation point indicated that these sentences were not perceived as ambiguous by participants before syntactic disambiguation took place. As the sentences were phonemically identical up to that point, some feature of the speech signal must have served a disambiguating role. The only possible cue for disambiguation prior to the disambiguation point was the acoustic information carried in the target word. Therefore, these findings suggest that the acoustic differences between noun and verb uses of noun/verb polysemes support the resolution of the syntactic ambiguity that these words could create in online language processing.

These findings can be integrated with previous research in which isolated noun and verb uses of noun/verb polysemes produced different neural responses [9]. The differences in the neural response to isolated exemplars suggested that brain had processed the phonetic differences. Likewise, neural differences have also been reported in response to noun and verb uses of polysemes in fluent speech [6–8], but that work did not consider whether those differences were the result of disambiguation at the level of the word itself.

Somewhat surprising was the lack of an N400 relative to the disambiguation point for the different replacement condition. If participants had used the acoustics of the target word to determine its grammatical category, then the different condition should have resulted in a semantic violation at the disambiguation point. We found no evidence of this pattern. Therefore, the information that resulted in disambiguation prior to that point may not have been confined to the word itself. Factors such as the relative frequencies of the sentence structures used in this study may also have contributed to listeners' expectations of how the ambiguous sentences would resolve. Another possibility is that the phonetic differences between the noun and verb uses were a byproduct of intonational boundaries following certain word uses. Previous research found final syllable nucleus and coda lengthening in words preceding an intonational boundary [16], and nouns were more likely to be followed by an intonational boundary than verbs [17]. Therefore, nouns were more likely to be lengthened than verbs, but the sentences in which they were used may have contained additional prosodic cues that were not carried directly on the target words themselves. This would explain why we found a difference in negativity in the N400 time window following the target words between the same and different replacement conditions even though the sentences had not yet been syntactically differentiated. Instead, the prosody surrounding the target word may have interacted with the acoustic features of the target word to give rise to a local semantic violation, but one that actually reduced ambiguity over the whole sentence. The present study cannot directly test this explanation because we neither manipulated nor assessed intonational boundaries in our stimuli. A follow up study that directly addresses intonational boundary placement may be better able to assess the role of such boundaries in ambiguity resolution in fluent speech.

The results of this study indicated that the low-level acoustic features that distinguished noun and verb uses of noun/verb polysemes played a role in ambiguity resolution in fluent speech. However, those may not have been the sole cue to the intended category of an ambiguous word. Rather, these differences appeared to interact with other aspects of sentence prosody, as well as structure frequency, to give rise to rapid, accurate interpretation

of fluent speech. These features work together to prevent the pervasive ambiguity in natural language from impeding comprehension.

Acknowledgements

This research was supported by pilot funding from grant 5P30 GM114748 from the National Institute of General Medical Sciences, part of the National Institutes of Health (Mark McCourt, PI). The contents of this article are the sole responsibility of the authors and do not reflect the official views of NIGMS or NIH. We also thank Ben Balas, Ganesh Padmanabhan, Allyson Saville, Amanda Auen, Rachel Helgeson, and Shaun Anderson for their assistance with study set-up, data collection, and data analysis.

Research reported here was funded by an Institutional Development Award from the National Institute of General Medical Sciences under grant number 5P30 GM114748, Mark McCourt, PI.

References

- [1]. Conwell E (2017). Prosodic disambiguation of noun/verb homophones in child-directed speech. *Journal of Child Language*, 44, 734–751. [PubMed: 26932727]
- [2]. Lohmann A & Conwell E (under review). Phonetic effects of grammatical category: How category-specific prosodic phrasing and token frequency impact the acoustic realization of nouns and verbs.
- [3]. Shi R, & Moisan A (2008). Prosodic cues to noun and verb categories in infant-directed speech. In Chan H, Jacob H, & Kipia E (Eds.), *BUCLD 32: Proceedings of the 32th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadia Press.
- [4]. Sorensen JM, Cooper WE, & Paccia JM (1978). Speech timing of grammatical categories. *Cognition*, 6, 135–153. [PubMed: 679646]
- [5]. Kelly MH (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, 99, 349–364. [PubMed: 1594729]
- [6]. Brown WS, Marsh JT, & Smith JC (1973). Contextual meaning effects on speech-evoked potentials. *Behavioral Biology*, 9, 755–761. [PubMed: 4764256]
- [7]. Brown WS, Marsh JT, & Smith JC (1976). Evoked potential waveform differences produced by the perception of different meanings of an ambiguous phrase. *Electroencephalography and Clinical Neurophysiology*, 41, 113–123. [PubMed: 58774]
- [8]. Brown WS, Marsh JT, & Smith JC (1979). Principal component analysis of ERP differences related to the meaning of an ambiguous word. *Electroencephalography and Clinical Neurophysiology*, 46, 709–714. [PubMed: 87317]
- [9]. Conwell E (2015). Neural responses to category ambiguous words. *Neuropsychologia*, 69, 85–92. [PubMed: 25637057]
- [10]. Kutas M, & Federmeier KD (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621–647.
- [11]. Frisch S, Schlesewsky M, Saddy D, & Alperman A (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, 85, B83–B92. [PubMed: 12169415]
- [12]. Friederici AD (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, 6, 78–84. [PubMed: 15866191]
- [13]. R Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing Retrieved from <http://www.R-project.org/>
- [14]. Bates D, Maechler M, Bolker BM, & Walker S (2015). Fitting linear mixed-effects models using lme4. ArXiv e-print Retrieved from <http://arxiv.org/abs/1406.5823>
- [15]. Kuznetsova A, Brockhoff PB, & Christensen RHB (2015). Package ‘lmerTest’ Retrieved from <https://cran.rproject.org/web/packages/lmerTest/index.html>
- [16]. Wightman CW, Shattuck-Hufnagel S, Ostendorf M, & Price PJ (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *The Journal of the Acoustical Society of America*, 91, 1707–1717. [PubMed: 1564206]

- [17]. Watson D, Breen M, & Gibson E (2006). The role of syntactic obligatoriness in the production of intonational boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 1045.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

ERP grand averages taken from central anterior sites (electrodes Fz, AFz, F1, and F2) during the post-target word window for all five conditions. Regions of analysis for the N400 and P600 are indicated.

Figure 2.

ERP grand averages taken from central anterior sites (electrodes Fz, AFz, F1, and F2) during the post-disambiguation window for all five conditions. Regions of analysis for the N400 and P600 are indicated.

Table 1.

Mixed model results for N400 window following target word offset

	Estimate (β)	Standard Error	df	t	p
Intercept	0.558	0.26	58.75	2.145	0.036
Grammatical Class	0.576	0.313	54	1.84	0.071
Replacement Type	0.655	0.313	54	2.091	0.041
Grammatical Class x Replacement Type	1.149	0.443	54	2.594	0.012

Table 2.

Mixed model results for P600 window following target word offset

	Estimate (β)	Standard Error	df	t	p
Intercept	1.072	0.322	55.28	3.329	0.002
Grammatical Class	0.561	0.376	54	1.493	0.141
Replacement Type	0.438	0.376	54	1.163	0.25
Grammatical Class x Replacement Type	1.114	0.532	54	2.094	0.041

Table 3.

Mixed model results for N400 window following disambiguating word offset

	Estimate (β)	Standard Error	df	t	p
Intercept	0.674	0.261	64.92	2.581	0.012
Grammatical Class	0.45	0.332	54	1.356	0.181
Replacement Type	0.405	0.332	54	1.22	0.228
Grammatical Class x Replacement Type	0.668	0.47	54	1.423	0.16

Table 4.

Mixed model results for N400 window following disambiguating word offset

	Estimate (β)	Standard Error	df	t	p
Intercept	0.741	0.3	65.75	2.391	0.02
Grammatical Class	0.582	0.397	54	1.466	0.148
Replacement Type	0.485	0.397	54	1.221	0.228
Grammatical Class x Replacement Type	0.687	0.562	54	1.224	0.226