Does acoustic unreliability of the speech signal affect N400?

Petra Hödl, Paul Iverson

Speech, Hearing and Phonetic Sciences, University College London (UCL), United Kingdom

p.hodl@ucl.ac.uk, p.iverson@ucl.ac.uk

Index Terms: intermittent noise, lexical processing, event-related potentials, N400

1. Introduction

Previous research suggests that listeners adjust their lexical processing to account for speech signal unreliability [1, 2]. Eye-tracking studies have shown that characteristics of lexical activation and competition during word recognition change under noise distortion [1] or when people listen to casual speech [2]. Under such adverse conditions lexical candidates that would be dismissed at early stages of word recognition remain activated for a longer period of time whereas other – normally stronger – candidates become weaker. Crucially, people not only delay their lexical decision when the actual input itself is distorted but also in situations in which they expect that the phonetic information is less reliable.

The aim of the current study was to investigate if adjustments of lexical processing to changes in the acoustic reliability of the speech signal can also be observed on a neurophysiological level. To address this question an event-related potential experiment was conducted, focusing on the N400 component. N400 was chosen because there is evidence supporting the view that N400 reflects activation of long-term memory representations in the lexicon – therefore lexical access processes [3, 4]. Importantly, it has also been shown that N400 is sensitive to stimulus degradation [5, 6, 7]. While in these experiments the speech input was degraded itself – either the final words [5], the sentence frames [6] or the whole sentences [7] – in the current study we also investigated whether N400 for completely undistorted speech changes when presented in the context of distorted speech – hence when the listening situation gives rise to question the acoustic reliability of the input.

2. Method

Our stimuli consisted of English sentences differing in the cloze probability of their final words (i.e., the probability that a particular word will be produced to complete a particular sentence). There were high cloze (e.g., "Monkeys like to eat yellow *bananas*"), low cloze (e.g., "It is very nice here in the *autumn*") and anomalous sentences (e.g., "You wash your hair using *lettuce*"). All sentences were recorded by a female native speaker of British English.

In order to manipulate acoustic reliability, we divided the material into two types of blocks: a baseline only containing undistorted stimuli and another with several conditions of stimuli with added intermittent noise that also included undistorted stimuli.

Stimulus distortion was achieved by adding short bursts of speech-shaped noise (signal-to-noise ratio of -3 dB SPL) to the signal. Those bursts varied in frequency and position within the sentences, and whether they were binaurally presented or not. In total there were three different noise conditions: "easy" noise in the sentence frame, "hard" noise in the sentence frame, and noise distortion only in the final word.

Exact positions of the noise bursts within the sentence frames were randomized as well as the sentence types. Crucially, in the mixed blocks there was no clue for the listeners to predict where the undistorted sentences would occur.

Subjects (n=11) were right-handed monolingual native speakers of British English. Their task was to listen carefully to the presented sentences and press a mouse button each time they hear an anomalous sentence.

Electroencephalograms were recorded from 64 scalp electrodes at a sampling rate of 2048 Hz using the Biosemi ActiveTwo EEG system. Testing consisted of a single experimental session and the material was divided in 10 blocks (2 baseline blocks and 8 noise blocks) of approximately 6 minutes with short breaks between each block. The 2 baseline blocks were either played in the beginning, in the middle or in the end of the experiment and before these blocks subjects were informed that they would be listening to sentences without noise.

3. Results and Discussion

Our preliminary results suggest that the N400 component is modulated by the perceived acoustic unreliability of the signal, even for stimuli with no added noise. That is, undistorted sentences in the baseline undistorted blocks produced a classic N400 effect (i.e., less N400 for high cloze sentences, and more for low cloze and anomalous sentences). However, undistorted high cloze sentences, when presented in the context of distorted sentences, appear to have a greater N400. That is, the sentences themselves are highly predictable (e.g., "Monkeys like to eat yellow *bananas*") so less lexical processing is needed, but in an unpredictable context listeners appear to delay their lexical decisions, considering more of the phonetic information and searching through the lexical competitors longer. This is in accord with previous eye-tracking studies [1, 2].

The preliminary results for the other intermittent noise conditions similarly indicate that masking the sentence context or the final word differentially affect the N400, in combination with the cloze probability of the sentences.

It thus appears that listeners modulate their lexical processing for speech to adapt to the listening conditions, both when there is actual masking of the stimulus and when there is an expectation of masking.

4. References

[1] J. M. McQueen and F. Huettig, "Changing only the probability that spoken words will be distorted changes how they are recognized," *Journal of the Acoustic Society of America*, vol. 131, no. 1, pp. 509–517, 2012.

[2] S. Brouwer, H. Mitterer, and F. Huettig, "Speech reductions change the dynamics of competition during spoken word recognition," *Language and Cognitive Processes*, vol. 27, no. 4, pp. 539–571, 2012.

[3] E. F. Lau, C. Phillips, and D. Poeppel, "A cortical network for semantics: (de)constructing the N400," *Nature Reviews Neuroscience*, vol. 9 no. 12, pp. 920–933, 2008.

[4] M. Kutas and K. D. Federmeier, "Electrophysiology reveals semantic memory use in language comprehension," *Trends in Cognitive Sciences*, vol. 4, no. 12, pp. 463–470, 2000.

[5] J. Aydelott, F. Dick, and D. L. Mills, "Effects of acoustic distortion and semantic context on event-related potentials to spoken words," *Psychophysiology*, vol. 43, no. 5, pp. 454–464, 2006.

[6] V. Boulenger, M. Hoen, C. Jacquier, and F. Meunier, "Interplay between acoustic/phonetic and semantic processes during spoken sentence comprehension: an ERP study," *Brain and Language*, vol. 116, no. 2, pp. 51–63, 2011.

[7] A. Strauß, S. A. Kotz, and J. Obleser, "Narrowed expectancies under degraded speech: revisiting the N400", *Journal of Cognitive Neuroscience*, vol. 25, no. 8, pp. 1383–1395, 2013.