Strategies for Longitudinal Neurophysiology. Commentary on Osterhout et al.

Doug Davidson
F. C. Donders Centre for Cognitive Neuroimaging
and
Max Planck Institute for Psycholinguistics

Osterhout et al. outline a research strategy to better understand the neurophysiological correlates of second language (L2) comprehension. Previous work has often employed cross-sectional experimental designs and group comparisons (e.g., the first language [L1] of native speakers vs. the L2 of bilinguals), and this work has shown variability in the response patterns in bilinguals. To reduce this variability, Osterhout et al. suggest an alternative research strategy in which novice learners are followed over time as they learn an L2. They argue that the common starting point of the novice learners will reduce the extra variability between individuals that would otherwise exist due to their prior L2 knowledge and skill and, furthermore, that by tracking individuals' rate of learning over time, subjects can serve as their own controls. This should then allow for greater experimental sensitivity to L1-L2 differences as L2 proficiency develops. They provide evidence from two series of studies: one concerning lexico-semantic acquisition and another concerning semantic and grammatical sentence processing using this type of design.

The second part of Osterhout et al.'s recommendation is to employ event-related potential (ERP) responses to study L2

Correspondence concerning this article should be addressed to Doug Davidson, F. C. Donders Centre for Cognitive Neuroimaging, P. O. Box 9101, 6500 HB Nijmegen, The Netherlands. Internet: Doug.Davidson@fcdonders.ru.nl

acquisition and comprehension. They suggest that ERP responses can reveal real-time sensitivity to semantic contrasts, as demonstrated by the N400 effect, and grammatical contrasts, via the P600 effect. It is important to emphasize the utility of this technique for understanding the neurophysiology of sentence processing. Much linguistic information, whether L1 or L2, appears in the form of sentences or clauses, and because words appear sequentially over time, a response measure is required that will show a response to each individual word as it is encountered. EEG or MEG measurements have the required temporal resolution for this measurement. As Osterhout et al. point out, other investigations of L2 processing using techniques such as functional magnetic resonance imaging (fMRI) have coarser temporal resolution and are thus less well suited for addressing questions about the time course of responses. Nevertheless, there are important limitations with electrophysiological techniques as well; for example, active areas that are located in the deep cortex or that produce closed fields, such as nuclei, might be difficult to observe using techniques such as EEG or MEG. Some evidence from language learning tasks using fMRI has shown involvement of thalamic nuclei (Opitz & Friederici, 2003). Especially for experimental designs in which changes are expected to occur over an extended period of time, it might be advantageous to have participants perform the same task in both fMRI and EEG to understand how changes in activity that is registered in fMRI are linked to changes observed with electrophysiological techniques (Dale & Halgren, 2001).

Osterhout et al. observed striking changes in the ERP response patterns in both the lexical and sentence processing tasks. In the first sentence processing study reported by Osterhout et al., however, they noted that variability between subjects remained as a dominant factor (a point they raise in the Discussion). Despite employing novice learners from the same university class and using textbook materials obtained from the classroom, the initial change in the response to grammatical violations was observed in only a subset of subjects. However, it should be emphasized that a group of subjects such as a university class is already quite homogenous. Students in a classroom setting are likely to have similar educational backgrounds with relatively high achievement in order to enroll in a university. The students would have the same instructor, the same homework assignments, and the same textbook. Thus, these subjects might be more homogenous, in terms of their educational background, prior language history, and classroom experience, than other groups of learners. Despite this, between-subject variability remained an important factor. Thus, it appears that other methods for assessing the impact of individual differences would be helpful, in conjunction with longitudinal designs. One possible approach would be to collect data from a group of tasks that measure a range of cognitive abilities and to use these as covariate measures for the changes in response patterns that are observed. Although this method does not eliminate individual differences, it could allow their influence to be modeled.

Finally, longitudinal experimental designs using students from a classroom setting could be used to examine learning on several different timescales. From an experimental point of view, there are many potential advantages to examining how responses change in students in a classroom setting. For example, the learning materials and vocabulary of the students can be used to construct experimental materials. Detailed information about teaching methods might be available, as well as information from a syllabus about when topics occur and when exercises and drills are performed. In addition, many behavioral and ERP experimental tasks share some similarity to the exercises in a classroom setting. In certain respects, an ERP experiment could be designed to be similar to a grammar learning exercise, as has already been done in some artificial grammar learning experiments (Opitz & Friederici, 2003). Thus, with this type of design, it might be possible to link short-term changes that occur within the time period of a single experiment (roughly corresponding to a classroom exercise) to the lasting longer-term changes in responses that occur over the duration of an entire course; for example, research on word repetition has shown that delayed repetition can lead to later responses than observed during immediate repetition (Rugg & Nagy, 1989) and that successful delayed recognition can lead to earlier responses than unsuccessful delayed recognition (Dhond, Witzel, Dale, & Halgren, 2005).

References

Dale, A. M., & Halgren, E. (2001). Spatiotemporal mapping of brain activity by integration of multiple imaging modalities. *Current Opinion in Neurobiology*, *11*, 202–208.

Dhond, R. P., Witzel, T., Dale, A. M., & Halgren, E. (2005). Spatiotemporal brain maps of delayed word repetition and recognition. *NeuroImage*, *28*, 293–304.

Opitz, B., & Friederici, A. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *NeuroImage*, *19*, 1730–1737.

Rugg, M. D., & Nagy, M. E. (1989). Event-related potentials and recognition memory for words. *Electroencephalographic and Clinical Neurophysiology*, *72*, 395–406.