Brain & Language 121 (2012) 12–24

Contents lists available at SciVerse ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Dissociating linguistic and non-linguistic gesture processing: Electrophysiological evidence from American Sign Language

Michael Grosvald [a,*], Eva Gutierrez [b], Sarah Hafer [b], David Corina [b,c]

[a] Department of Neurology, University of California at Irvine, United States
[b] Center for Mind and Brain, University of California at Davis, United States
[c] Departments of Linguistics and Psychology, University of California at Davis, United States

ARTICLE INFO

Article history:
Accepted 24 January 2012
Available online 17 February 2012

Keywords:
Sign language
ASL
ERP
N400
Deaf
Pseudo-word
Grooming gesture

ABSTRACT

A fundamental advance in our understanding of human language would come from a detailed account of how non-linguistic and linguistic manual actions are differentiated in real time by language users. To explore this issue, we targeted the N400, an ERP component known to be sensitive to semantic context. Deaf signers saw 120 American Sign Language sentences, each consisting of a "frame" (a sentence without the last word; e.g. BOY SLEEP IN HIS) followed by a "last item" belonging to one of four categories: a high-close-probability sign (a "semantically reasonable" completion to the sentence; e.g. BED), a low-close-probability sign (a real sign that is nonetheless a "semantically odd" completion to the sentence; e.g. LEMON), a pseudo-sign (phonologically legal but non-lexical form), or a non-linguistic grooming gesture (e.g. the performer scratching her face). We found significant N400-like responses in the incongruent and pseudo-sign contexts, while the gestures elicited a large positivity.

1. Introduction

While it is now widely accepted that signed languages used in deaf communities around the world represent full-fledged instantiations of human languages—languages which are expressed in the visual–manual modality rather than the aural–oral modality—the question of how a sign is recognized and integrated into a sentential context in real time has received far less attention (see Corina & Knapp, 2006; Emmorey, 2002; for some discussions). Sign language recognition may be more complicated than spoken language recognition by virtue of the fact that the primary articulators, the hands and arms, are also used in a wide range of other common everyday behaviors that include non-linguistic actions such a reaching and grasping, waving, and scratching oneself, as well gesticulations that accompany speech (i.e. co-speech gestures) or serve non-sign language deictic functions, such as pointing.

The formal relationship between signed languages and human gestural actions is of considerable interest to a range of disciplines. Linguists, psychologists and cognitive scientists have proposed a critical role for manual gesture in the development and evolution of human languages (Arbib, 2005, 2008; Gentilucci & Corballis, 2006; Rizzolatti & Arbib, 1998; Tomasello, 2005; Wilcox, 2004). Recently, linguists have documented compelling evidence that the development of nascent sign languages derives from idiosyncratic gestural and pantomimic systems used by isolated communities, which in some cases may be limited to individual families who have a need to communicate with a deaf child (Frishberg, 1987; Goldin-Meadow, 2003; Kegl, Senghas, & Coppola, 1999; Meir, Sandler, Padden, & Aronoff, 2010; Morford & Kegl, 2000; Senghas, 2005). Even within mature sign languages of Deaf communities, linguistic accounts of sign language structure have also argued that lexical and discourse components of American Sign Language (ASL) and other signed languages may be best understood as being gesturally based (Liddell, 2003). Thus diachronic and synchronic evidence from language research support the contention that signed languages might make use of perceptual systems similar to those through which humans understand or parse human actions and gestures more generally (Corballis, 2009). In contrast, given its linguistic status, sign language perception may require the attunement of specialized systems for recognizing sign forms.

A comprehensive theory of sign language recognition will be enhanced by providing an account of when and how the processing of sign forms diverges from the processing of human actions in general. Recent behavioral and neuro-imaging studies have reported differences in deaf subjects' responses to single signs compared to non-linguistic gestures (Corina, Grosvald, & Lachaud, 2011; Corina et al., 2007; Emmorey, Xu, Gannon, Goldin-Meadow, & Braun, 2010; MacSweeney et al., 2004), but no studies to our knowledge have examined the recognition of signs and gestures

* Corresponding author. Address: Department of Neurology, University of California at Irvine, 101 The City Drive South, Bldg. 53, Room 204, Orange, CA 92868-4280, United States. Fax: +1 714 456 1697.
E-mail address: m.grosvald@uci.edu (M. Grosvald).

0093-934X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2012.01.005

under sentence processing constraints. Consider for example the signer, who, in mid-sentence, fulfills the urge to scratch his face, or perhaps swat away a flying insect. What is the fate of this non-linguistic articulation? Does the sign perceiver attempt to incorporate these manual behaviors into accruing sentential representations, or are these actions easily tagged as non-linguistic and thus rejected by the parser? The goal of the present paper was to use real-time electrophysiological measures to assess empirically the time course of sentence processing in cases where subjects encountered non-linguistic manual forms (here "self-grooming" behaviors, e.g. scratching the face, rubbing one's eye, adjusting the sleeves of a shirt, etc.). We sought to compare the processing of these non-linguistic gestural forms within a sentential context to cases in which deaf signers encountered violations of semantic expectancy that have been observed to elicit a well-defined electrophysiological component, the N400.

The N400 component (Holcomb & Neville, 1991; Kutas & Hillyard, 1980) has been frequently investigated in previous ERP research on written, spoken and signed language (e.g. Capek et al., 2009; Kutas, Neville, & Holcomb, 1987). The N400 is a broad negative deflection generally seen at central and parietal scalp sites that peaks about 400 ms after the visual or auditory presentation of a word. Although all content words elicit an N400 component, the ERP response is larger for words that are semantically anomalous or less expected (Hagoort & Brown, 1994; Kutas & Hillyard, 1984); thus the N400 is often interpreted as an index of ease or difficulty in semantic conceptual integration (Brown & Hagoort, 1993; Hagoort & Van Berkum, 2007). For example, for listeners encountering the two sentences "I like my coffee with milk and *sugar*" and "I like my coffee with milk and *mud*," the N400 response to the last word in the second item is expected to be larger.

An N400 or N400-like component can also be found in response to orthographically/phonologically legal but non-occurring "pseudo-words" (e.g. "*blork*"), and it has sometimes been reported that pseudo-words elicit a stronger N400 response than semantically incongruent real words (Bentin, 1987; Bentin, McCarthy, & Wood, 1985; Hagoort & Kutas, 1995), consistent with the idea that the magnitude of N400 response is related to the difficulty of the ongoing process of semantic-contextual integration. However, orthographically *illegal* "non-words" (e.g. "*rbsnk*") do not generally elicit an N400, and a *positive* component is sometimes seen instead (Hagoort & Kutas, 1995; Ziegler, Besson, Jacobs, Nazir, & Carr, 1997). This may reflect the operation of some kind of filtering mechanism during online processing, through which language users are able to quickly reject forms that lie beyond a certain point of acceptability, or plausibility, during the ongoing processing of the incoming language stream.[1]

The N400 (or N400-like responses) can also be observed in numerous contexts involving non-linguistic but meaningful stimuli, such as pictures (Ganis & Kutas, 2003; Ganis, Kutas, & Sereno, 1996; Nigam, Hoffman, & Simons, 1992; Pratarelli, 1994), faces (Barrett & Rugg, 1989; Bobes, Valdés-Sosa, & Olivares, 1994), environmental noises (Chao, Nielsen-Bohlman, & Knight, 1995; Van Petten & Rheinfelder, 1995), movie clips (Sitnikova, Holcomb, Kiyonaga, & Kuperberg, 2008; Sitnikova, Kuperberg, & Holcomb, 2003) and co-speech gestures (Kelly, Kravitz, & Hopkins, 2004; Wu & Coulson, 2005).

Linguistically anomalous stimuli are not always associated with an N400 response. For example, the left anterior negativity (LAN; Friederici, 2002; Neville, Nicol, Barss, Forster, & Garrett, 1991) and P600 (Osterhout & Holcomb, 1992) are well-known ERP components that have been found in syntactic violation contexts in spoken and written language, and more recent work has shown that these components can be elicited in the visual–manual modality as well. For example, in a recent study Capek et al. (2009) compared ERP responses to semantically and syntactically well-formed and ill-formed sentences. While semantic violations elicited an N400 that was largest over central and posterior sites, syntactic violations elicited an anterior negativity followed by a widely distributed P600. These findings are consistent with the idea that within written, spoken and signed languages, semantic and syntactic processes are mediated by non-identical brain systems (Capek et al., 2009).

The present study makes use of dynamic video stimuli showing ASL sentences completed by four classes of ending item—semantically congruent signs, semantically incongruent signs, phonologically legal but non-occurring pseudo-signs, and non-linguistic grooming gestures. Based upon previous studies, we expected a gradation of N400-like responses across conditions, with N400 effects of smaller magnitude for semantically incongruent endings and of larger magnitude (i.e. more negative) for phonologically legal pseudo-signs.

The ERP response for the non-linguistic gesture condition is *a priori* more difficult to predict. Previous neuro-imaging studies of deaf signers have reported differences in patterns of activation associated with the perception of signs compared to non-linguistic gestures (Corina et al., 2007; Emmorey et al., 2010; MacSweeney et al., 2004), but the methodologies used in those studies lacked the temporal resolution to determine at what stage of processing these differences may occur. While N400-like responses have been elicited to co-speech gestural mismatches (Kelly et al., 2004; Wu & Coulson, 2005), in our study, gestures occur *in place of* semantically appropriate sentence-ending items, rather than as a possible accompaniment. It should also be borne in mind that the relationship of signs and grooming gestures is probably not quite akin to that between standard lexical items in spoken language and the orthographically/phonotactically illegal pseudo-words used in earlier ERP studies. Unlike grooming gestures, which are part of everyday life, illegal non-words like "dkfpst" are probably alien to most people's routine experience. A better spoken-language analogue of our grooming action condition might be something like "I like my coffee with milk and *[clearing of throat]*," though we know of no spoken-language studies which have incorporated such a condition. The non-linguistic grooming gestures used in the present study may be another example of forms that language users (in this case, signers) are able to quickly reject as non-linguistic during language processing. If this is the case, then one might also expect that such forms will *not* elicit an N400 but rather a positive-going component (cf. Hagoort & Kutas, 1995).

In summary, to the extent that semantic processing at the sentence level is similar for signed and spoken language, despite the obvious difference in modality, the ERP responses associated with our four sentence ending condition should be predictable. First, the incongruent signs should elicit a negative-going component relative to the baseline (congruent sign) condition, consistent with the classic N400 response seen for English and other spoken languages, as well as some previous ERP studies of ASL (Kutas et al., 1987; Neville et al., 1997). Second, the pseudo-signs should also elicit a negative-going wave, and this response can be expected to be of larger magnitude (i.e. be more negative) than that seen for the incongruent signs. Third, while the likely response to the grooming gesture condition is more difficult to predict, we may expect to see a positive-going component relative to the baseline.

2. Methodology

2.1. Participants

The 16 participants (12 female and 4 male; age range = [19, 45], mean = 25.4 and SD = 8.3) were deaf users of ASL; all were students

[1] This possibility is bolstered by recent work of Albert Kim and colleagues, who have found that relative to real word controls, N400 amplitude decreases and P600 amplitude increases, parametrically, as orthographic irregularity increases (Kim & Pitkänen, submitted for publication).