Cognition 166 (2017) 433–446

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Original Articles

Predicting semantic features in Chinese: Evidence from ERPs

Nayoung Kwon [a,*], Patrick Sturt [b], Pan Liu [c]

[a] Department of English, Konkuk University, Republic of Korea
[b] Department of Psychology, University of Edinburgh, UK
[c] Behavioural Sciences Institute, Singapore Management University, Singapore

ARTICLE INFO

Article history:
Received 22 June 2016
Revised 13 March 2017
Accepted 8 June 2017
Available online 20 June 2017

Keywords:
Prediction
Classifiers
Chinese
ERPs
N400
Sentence processing

ABSTRACT

This article reports two ERP studies that exploited the classifier system of Mandarin Chinese to investigate semantic prediction. In Mandarin, in certain contexts, a noun has to be preceded by a classifier, which has to match the noun in semantically-defined features. In both experiments, an N400 effect was elicited in response to a classifier that mismatched an up-coming predictable noun, relative to a matching classifier. Among the mismatching classifiers, the N400 effect was graded, being smaller for classifiers that were semantically related to the predicted word, relative to classifiers that were semantically unrelated to the predicted word. Given that the classifier occurred before the predicted word, this result shows that fine-grained semantic features of nouns can be pre-activated in advance of bottom-up input. The studies thus extend previous findings based on a more restricted range of highly grammaticalized features such as gender or animacy in Indo-European languages (Szewczyk & Schriefers, 2013; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Bates, Moreno, & Kutas, 2003).

1. Introduction

Prediction or anticipation refers to a mental process that generates information about future states based on what we know already. As one of the most fundamental principles of human cognition (Clark, 2013), prediction operates in various cognitive domains, such as visual processing, motor control and theory of mind (Friston & Stephan, 2007; Frith & Frith, 2006; Mehta & Schaal, 2002; Wolpert & Flanagan, 2001), and its importance has also been emphasized in studies of language processing (DeLong, Urbach, & Kutas, 2005; Federmeier, 2007; Pickering & Garrod, 2007; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). However, although there has been some compelling evidence for semantic prediction, most of the relevant studies were conducted based on unilateral semantic distinctions such as gender or animacy and in Indo-European languages, in which semantic features such as gender and/or animacy are highly grammaticalized, and correlate with overt morpho-syntactic markers. To evaluate the generality of semantic prediction, in this study we use the classifier system of Mandarin Chinese, in which agreement between a classifier and a noun is based on perceived similarity that is functionally or perceptually defined, representing various semantic relations.

1.1. Prediction in language processing

Language processing has been argued to be highly incremental and predictive. Previous studies have shown that language comprehenders do not delay the processing of incoming words until the end of a sentence, despite temporary structural ambiguity rampant in natural human language (Frazier & K. Rayner, 1988). Instead, the incoming words are parsed and interpreted immediately as they are perceived. This incremental nature of language processing is evidenced by garden-path effects, in which processing difficulty is elicited at the point of structural disambiguation when the initially built structure turns out to be incorrect and thus requires structural revision (Rayner & Frazier, 1987; Frazier & Rayner, 1982). On the other hand, the predictive nature of language processing has been mainly discussed in association with context effects. That is, it is argued that contextual information facilitates processing of a target item because that item has been predicted ahead of time based on the preceding context, as evidenced by shorter reaction times in a simple reading or lexical judgment task (Ehrlich & Rayner, 1981; Fischler & Bloom, 1979; Jordan & Thomas, 2002; Hess, Foss, & Carroll, 1995; Wright & Garrett, 1984) or by predictive eye-gaze patterns in the visual-world paradigm. For example, using the visual-world paradigm, Kamide, Altmann, and

* Corresponding author at: Department of English, School of Humanities, Konkuk University, Neungdong-ro 120, Gwangjin-gu, Seoul 143-701 Republic of Korea.
E-mail address: nayoung.kw@gmail.com (N. Kwon).

http://dx.doi.org/10.1016/j.cognition.2017.06.010
0010-0277/

Haywood (2003) showed that sentence comprehenders moved their eyes towards a picture of an object that was appropriate given the context—for example, there were more fixations on a picture of a motorbike than a carousel given the context *The man will ride,* but the reversed pattern was observed given the context *The girl will ride.* Given that these effects were already significant at the verb position even before the participants perceived the object in the speech stream, Kamide at al. took these results to suggest that sentence comprehenders predict the semantic contents of an upcoming word based on the combinatory semantic constraints of an agent (*the man*) and a verb (*ride*) (for similar results, see also Kamide, Scheepers, & Altmann, 2003; cf. Kaiser & Trueswell, 2004; Weber, Grice, & Crocker, 2006).

Recent studies have suggested that ERPs also provide efficient means to examine the effects of prediction during on-line sentence processing (DeLong et al., 2005; Van Petten & Luka, 2012; Lau, Stroud, Plesch, & Phillips, 2006; and references therein). In particular, the N400 component has proven to be useful in the investigation of prediction effects in language processing. The N400 is a negative-going wave that peaks about 400 ms post-stimulus, with a centro-parietal maximum. It was first observed in response to semantic anomalies such as *He spread the warm bread with socks* (Kutas & Hillyard, 1980) and has been since shown to be sensitive to the processing of semantic information regardless of semantic anomaly (see Kutas & Federmeier, 2011 for a review). For example, the amplitude of N400 varies depending on an item's predictability in a given context (cf. Chow et al. 2014). Thus, N400 amplitude increases in response to a less probable sentential continuation (e.g. *dog* for *Don't touch the wet* ...) relative to the expected one (e.g. *paint*) (Kutas & Hillyard, 1984). In addition, the amplitude of the N400 also varies as a function of frequency of an item or as a function of semantic relations independent of the local context. Thus, a high frequency word elicits a smaller N400 compared with a low frequency a word, when they are presented in isolation (Rugg, 1990). Similarly, a high typicality category member (e.g. robin as a type of bird) elicits a smaller N400 compared to a low typicality category member (e.g. turkey as a type of bird), which in turn elicits a smaller N400 compared to an unrelated item (Kutas & Federmeier, 2000).

More relevant to the goal of this paper is the study of Federmeier and Kutas (1999). In this study, the authors compared ERP responses to two types of semantic category violation: 'within-category violation', with an unexpected item from the same semantic category as the predicted item, and 'between-category violation' with an equally unexpected item from a different semantic category. For example, when preceded by a lead-in sentence like *They wanted to make the hotel look more like a tropical resort*, a second target sentence (*So along the driveway, they planted rows of* ...) ended either with (i) the expected item (e.g. *palms*), (ii) a within-category violation (e.g. *pines*), or (iii) a between-category violation (e.g., *tulips*). The results showed that both the two unexpected endings elicited larger N400s than the predicted ending. Importantly, however, the between-category violation elicited still larger N400s than the within-category violation, despite their equally low cloze probability (see also Federmeier, McLennan, De Ochoa, & Kutas, 2002; Kutas & Hillyard, 1984; Kutas, 1993; Thornhill & Van Petten, 2012 for similar results). These results suggest that the neural representation of a word is based on a set of semantic features, and that an overlap in semantic features facilitates the processing of an unexpected item in on-line language processing. In the context of the current study, Federmeier and Kutas' (1999) results are compatible with the claim that some semantic content was predicted ahead of time. However, the results are also compatible with the integration hypothesis, namely that within-category violations were easier to integrate with the preceding context due to more overlap in semantic features than between-category violations (cf. Wlotko & Federmeier, 2015).

On the other hand, DeLong et al. (2005) provides stronger evidence for prediction effects. Using the phonological regularity of the English indefinite articles *a* and *an*, DeLong et al. examined ERP responses to sentence contexts that led to anticipation of specific words that started either with a consonant or a vowel. For example, *The day was breezy so the boy went outside to fly* ... was followed by *a kite*, a continuation that is highly expected, or *an airplane*, a continuation that is plausible but less likely. The study replicated the well-known correlation between N400 amplitude and cloze probability of target nouns (Kutas & Hillyard, 1984) with bigger N400 responses to *airplane* than to *kite.* More importantly, the strong correlation was also evident at the preceding article, with bigger N400 responses to *an* (an article preceding an unexpected noun *airplane*) compared to *a* (an article preceding an expected noun *kite*). Given that phonological variations of the two indefinite articles are based solely on the word forms of the predicted target nouns, the results at the article strongly suggest that sentence comprehenders predict specific words or specific phonological forms of words (for further evidence of form-based expectations, see DeLong, Groppe, Urbach, & Kutas, 2012; Ito, Corley, Pickering, Martin, & Nieuwland, 2016; Kim & Lai, 2012; Laszlo & Federmeier, 2009; Martin et al., 2013).

Similar results were also obtained from studies of Spanish, in which an article should agree with a noun in gender (e.g. *una canasta* 'a$_{_feminine}$ basket$_{_feminine}$'). Capitalizing on this grammatical gender system, Wicha, Bates, Moreno, and Kutas (2003) (see also Wicha, Moreno, & Kutas, 2003) showed that an unexpected article (e.g. *un* 'a$_{_masculine}$') elicited broadly distributed N400-like negativity compared to an expected article (e.g. *una* 'a$_{_feminine}$'). As the gender of an article is determined by the following noun (e.g. *canasta* 'basket$_{_feminine}$'), which is, in turn, predicted based on the prior sentence context (e.g. 'Red riding Hood carried the food for her grandmother in ∼'), the authors took these N400 effects as evidence that readers predict gender-specific nouns (see also Szewczyk & Schriefers, 2013 for similar experimental results based on animacy manipulations in Polish).

A study in Dutch with a similar experimental design further supported the relevance of prediction in language processing, but the prediction effect in this study was observed in the form of positivity (Van Berkum et al., 2005). Similarly to Spanish, Dutch also has a grammatical gender system in which the gender of a preceding adjective has to agree with that of a noun, in indefinite noun phrases. Van Berkum et al. (2005) found that after a highly constraining discourse context (e.g. 'The burglar had no trouble locating the secret family safe. Of course, it was situated behind a...'), adjectives (e.g. *grote* 'big$_{_common}$') whose gender mismatched with that of the predicted noun (e.g. *schilderij* 'painting$_{_neutral}$') elicited a larger fronto-central positivity compared with their gender-matched counterparts (e.g. *groot* 'big$_{_neutral}$') (see also Wicha, Moreno, & Kutas, 2004 for positive-going effects with gender manipulations in Spanish).

Van Petten and Luka (2012) broadly characterized the frontal positivity as indexing costs of failed prediction, while the N400 effect as related to "the other side of coin", an index of processing facilitation due to successful prediction. However, it is not clear at this point why highly analogous studies have elicited different ERP responses. That is, monophasic N400 and frontal positivity effects have each been independently elicited in studies of a similar design (negativity: Otten & Van Berkum, 2009; Wicha, Bates, et al., 2003; Wicha, Moreno, et al., 2003; positivity: Van Berkum et al., 2005; Wicha et al., 2004) while other studies have shown these two components simultaneously, in a bi-phasic effect (Delong, Urbach, Groppe, & Kutas, 2011; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007; Kutas, 1993; Otten, Nieuwland, & Van Berkum,