Brain Research

Research Report

Does the ending matter? The role of gender-to-ending consistency in sentence reading

Sendy Caffarra[a,*], Horacio A. Barber[a,b]

[a]BCBL, Basque Center on Cognition, Brain and Language, Paseo Mikeletegi 69, 20009 Donostia-San Sebastian, Spain
[b]Department of Cognitive Psychology, and Institute of Biomedical Technologies, University of La Laguna, Campus de Guajara, 38205 Tenerife, Spain

ARTICLE INFO

ABSTRACT

In many languages, during language comprehension the cognitive system needs to recover grammatical gender features in order to identify agreement dependencies established between different sentence constituents. A two-route model proposes that gender can be retrieved either lexically or computing its correlations with the word-form. However, evidence supporting this model has been collected thus far only with metalinguistic tasks on isolated nouns or word pairs. The present ERP study was aimed at testing whether the system is sensitive to gender formal cues within a sentence context. Specifically, we investigated the time course of gender processing in sentence reading where the target nouns could show a reliable gender-related ending (i.e., transparent nouns) or an ambiguous ending (i.e., opaque nouns). The results showed a greater central-anterior negativity for transparent nouns than for opaque nouns between 200 ms and 500 ms, suggesting that the system can rapidly detect reliable formal cues to gender. In addition, gender agreement violations showed a LAN-P600 pattern that was not modulated by the gender-to-ending consistency. Taken together, these results confirm that also during sentence comprehension, distributional gender cues conveyed by noun endings can be detected. This finding is compatible with the existence of a form-based route. The formal cues to gender are detected at an early stage, this probably being part of the word recognition process. Whereas this distributional information does not seem to be crucial in computing agreement dependencies within a sentence context.

1. Introduction

One-fourth of the world's languages have a grammatical gender system where nouns can be classified as belonging to different classes (e.g., feminine, masculine, neuter). This grammatical feature had been widely investigated due to its syntactic relevance. In fact, during sentence comprehension, the cognitive system needs to recover the gender value of different sentence constituents in order to identify agreement dependencies and built the sentence structure.

One of the main research questions on this issue has asked how gender information can be represented and

*Corresponding author.
    E-mail addresses: s.caffarra@bcbl.eu (S. Caffarra), hbarber@ull.es (H.A. Barber).

recovered. In the domain of language comprehension, one of the theoretical frameworks that has been proposed in order to describe gender retrieval (Gollan and Frost, 2001) claims that the gender value of a noun can be retrieved using two different routes, a lexical and a form-based route.

The lexical route derives gender value as an abstract feature stored in the mental lexicon. This way, each stored noun representation is associated with its corresponding gender information. Neuropsychological studies showed that the lexical route can always provide the correct gender of a noun, regardless of its phonological or orthographic features (Badecker et al., 1995; Miozzo and Caramazza, 1997; Vigliocco et al., 1997) and, for this reason, it is considered to be essential in recovering gender information (Delfitto and Zamparelli, 2009). However, it should be noted that, in a number of languages (e.g., Romance languages), gender can entail strong correlations with the formal features of the word and these consistent associations can represent an important distributional cue for gender retrieval (Corbett, 1991). In Spanish, for example, almost all the nouns that end in –o are masculine (there are only few exceptions to this rule, such as *mano*, "hand$_F$"), while the nouns ending in –a are feminine in 96% of the cases (Harris, 1991). Therefore, it would be surprising if the cognitive system did not take advantage of this additional source of information at the time of gender identification (and it would be surprising if languages kept these correlations through history if they lacked any cognitive function). Empirical evidence has shown, for example, that strong gender-to-ending consistency helps learn what gender class nouns belong to (Caselli et al., 1993; Matthews, 2010; Pérez-Pereira, 1991), and it can also be useful when system needs to recover gender information during language comprehension (Hernandez et al., 2004). Different theoretical frameworks have been proposed in order to account for the system's sensitivity to distributional features of a language (Saffran et al., 1996; MacWhinney et al., 1984). A potential explanation that has been proposed within the specific case of gender retrieval assumes the presence of a second form route, which would recover gender information on the basis of its correlations with the word-form level (Gollan and Frost, 2001).

According to this two-route model, gender can be differently retrieved depending on the availability of gender-related endings. Thus, if we consider the Spanish nouns *queso* ("cheese$_M$") and *reloj* ("watch$_M$"), the Gollan and Frost's hypothesis (2001) predicts that they should differ in the way in which gender can be retrieved. Specifically, in the case of *queso* ("cheese$_M$"), a gender-related ending is available and the system has the possibility to use both routes: it can recover gender as an abstract feature stored in the lexicon (lexical route), and it can take advantage of the distributional cues conveyed by the word ending (form-based route). In the case of the Spanish noun *reloj* ("watch$_M$"), the ending is uninformative of grammatical gender (i.e., the ending "j" is not consistently associated with a specific gender class) and the lexical route is the only way to provide the correct gender of the word.

The first group of nouns, whose ending has a regular correspondence with a specific gender class, are typically called transparent nouns; while the second group of nouns, whose ending is ambiguous, are called opaque nouns (Bates et al., 1995). According to the two-route hypothesis, in both cases the lexical route is supposed to be activated and the only difference between these two groups of nouns concerns the availability of the form-based route. This implies that the system would be able to detect any regular correspondence between the word-form and a specific gender class.

Several behavioural studies have tested whether the system would be sensitive to the presence of gender-related noun endings adopting a between-item comparison design where two groups of nouns were presented out of a sentence context (Bates et al., 1995, 1996; De Martino et al., 2010; Gollan and Frost, 2001; Hernandez et al., 2004; Padovani and Cacciari, 2003; Padovani et al., 2005; Peereman et al., 2009; Schiller et al., 2003; Taft and Meunier, 1998). The results showed that when transparent and opaque nouns were presented in isolation, participants were faster and/or more accurate in performing a lexical decision task (Padovani et al., 2005) or a gender decision task (Bates et al., 1995; Hernandez et al., 2004; Gollan and Frost, 2001). These findings suggest that the system is sensitive to the presence of gender-related endings and it can rely on these formal cues to access gender. The authors interpreted these results as suggesting that an additional form-based route facilitates gender retrieval in the case of transparent nouns (Gollan and Frost, 2001).

Nonetheless, it is still unclear under what circumstances this second form-based route can be activated. Behavioural studies on word pairs presented in isolation (e.g., noun–adjective) showed that, when subjects had to perform a gender decision task on a target noun or when they had to judge the gender agreement between two words, there was a behavioural advantage for transparent nouns (Bates et al., 1996). This effect was observed for both agreeing and disagreeing word pairs (Bates et al., 1996; cf. Gollan and Frost, 2001). Interestingly, when the participants were asked to simply repeat words, no difference was found between transparent and opaque nouns (Bates et al., 1995, 1996; Padovani and Cacciari, 2003). On the basis of these data authors proposed that formal cues to gender might be used at a late stage of checking that participants engaged especially when they were required to make an explicit decision on gender identity or agreement. According to the post-lexical checking hypothesis (Bates et al., 1996), the system would primarily rely on the lexical route. When conflicting information exists between the gender specified in the noun representation (the lexical route) and the noun ending (the form-based route), a post-lexical checking procedure would check for gender-marking regularity. Hence, the form-based route would intervene at a later stage of analysis to monitor the output of the lexical route, or to recover from gender agreement errors (Bates et al., 1995, 1996; Gollan and Frost, 2001).

Taken together, these experimental findings suggest that when the system needs to recover gender information, it detects the presence of formal cues to gender and takes advantage of this information. Although these behavioural findings seem to confirm Gollan and Frost's predictions (2001), it should be noted that this positive evidence comes from studies on isolated nouns or isolated word pairs. However, deciding gender identity or gender agreement on target words presented in isolation may require strategies that do not normally characterize gender retrieval (Radeau and Van Berkum, 1996). In addition, the comprehension of a complex sequence of words is characterized