Brain & Language 122 (2012) 71–80

Contents lists available at SciVerse ScienceDirect

Brain & Language

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

The neuronal infrastructure of speaking

Laura Menenti [a,b,*], Katrien Segaert [a,c], Peter Hagoort [a,c]

[a] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands
[b] Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, United Kingdom
[c] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Accepted 23 April 2012
Available online 19 June 2012

Keywords:
Language production
fMRI
fMRI adaptation
Repetition suppression
Semantics
Syntax
Lexical processing

ABSTRACT

Models of speaking distinguish producing meaning, words and syntax as three different linguistic components of speaking. Nevertheless, little is known about the brain's integrated neuronal infrastructure for speech production. We investigated semantic, lexical and syntactic aspects of speaking using fMRI. In a picture description task, we manipulated repetition of sentence meaning, words, and syntax separately. By investigating brain areas showing response adaptation to repetition of each of these sentence properties, we disentangle the neuronal infrastructure for these processes. We demonstrate that semantic, lexical and syntactic processes are carried out in partly overlapping and partly distinct brain networks and show that the classic left-hemispheric dominance for language is present for syntax but not semantics.

1. Introduction

Twas brillig and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves
And the mome raths outgrabe.(Carroll, 1871)

What are toves? No clue. Or, well…? They are things that can be plural, be 'slithy', and can 'gyre' and 'gimble'. We are apparently able to process grammar in the absence of meaningful words. Indeed, models of speech production (Garrett, 1975; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999; Vigliocco & Hartsuiker, 2002) agree that what we say (semantics) and how we say it (syntax) are different aspects of the speaking process. In this study, we investigate whether the neuronal infrastructure underlying speaking also shows this distinction.

Speaking is the conversion of an intention to communicate a message into a linearized string of speech sounds. An essential step in this process is the retrieval of the relevant concepts and the specification of semantic structure. One key aspect of semantics is *thematic role structure*. It refers to the relation between the different concepts and events, or 'who does what to whom'. Schematically one can state this as a predicate with arguments: BITE(DOG, MAN) means there is a BITE event, a DOG is the agent of this event (the one who bites), while MAN is the patient (the one who is bitten).

The thematic role structure BITE(DOG, MAN) can be expressed in a variety of ways, depending on the choice of a syntactic structure: "The man was bitten by a dog", "The dog bit a man", " It was a dog that bit the man", or even "Did the dog bite a man?". Of course, though these sentences share the same thematic role structure, they do have subtle differences in meaning. In this paper we will focus on thematic role structure as a key aspect of semantic structure. Words play a central role in connecting semantic and syntactic structure (Hagoort, 2005; Levelt, 1989; Vosse & Kempen, 2000). The mental lexicon, our memory for language, contains information on semantic, syntactic and phonological properties of words. When we prepare an utterance, the relevant concepts (TO BITE, DOG, MAN) are retrieved from memory. The syntactic properties (e.g., word class, grammatical gender) of the associated words are in turn also activated. For example, the concept BITE belongs to a lexical item 'to bite' that has the syntactic property of being a verb. This verb takes a subject and an object, which defines the specifics of the sentence structure that the verb can enter into. For example, the sentence "The dog bites a man to the woman" is ungrammatical because the verb 'bites' cannot take an indirect object. Through unification of these syntactic constraints of activated lexical items a syntactic structure for the sentence is generated.

* Corresponding author. Address: Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands.
E-mail address: laura.menenti@mpi.nl (L. Menenti).

0093-934X/$ - see front matter
http://dx.doi.org/10.1016/j.bandl.2012.04.012

Though cognitive models specify the components of speech production in much detail, very little is known about the neuronal infrastructure for the linguistic processes involved in producing multi-word utterances. So far cognitive neuroimaging research on language has focused on comprehension (Gernsbacher & Kashak, 2003), and the few studies on language production are mostly either on single word production (Alario, Chainay, Lehericy, & Cohen, 2006; Karbe, Herholz, Weber-Luxenburger, Ghaemi, & Heiss, 1998; Kircher, Brammer, Tous Andreu, Williams, & McGuire, 2001; Tremblay & Gracco, 2006; Tremblay & Gracco, 2010; Tremblay & Small, 2011b; Wise et al., 2001; Zheng, Munhall, & Johnsrude, 2010; for a review, see Indefrey & Levelt, 2004), or on covert production (den Ouden, Hoogduin, Stowe, & Bastiaanse, 2008). The neuroimaging studies that have investigated overt sentence-level production either treat sentence production as a unitary process (Awad, Warren, Scott, Turkheimer, & Wise, 2007; Blank, Scott, Murphy, Warburton, & Wise, 2002; Brownsett & Wise, 2010; Foki, Gartus, Gesissler, & Beisteiner, 2008; Kemeny, Ye, Birn, & Braun, 2005; Kircher, Brammer, Williams, & McGuire, 2000; Stephens, Silbert, & Hasson, 2010), or isolate only one component of speech production (Haller, Radue, Erb, Grodd, & Kircher, 2005; Indefrey et al., 2001; Kircher, Oh, Brammer, & McGuire, 2005; Tremblay & Small, 2011a). It is, therefore, unknown to what degree the different cognitive stages in speech production also recruit different neuronal networks.

In a recent functional Magnetic Resonance Imaging (fMRI) study, we have disentangled semantic, lexical, and syntactic processes during sentence production and comprehension by using a fMRI adaptation paradigm (Menenti, Gierhan, Segaert, & Hagoort, 2011). fMRI adaptation is a phenomenon in which the Blood Oxygen Level Dependent (BOLD) response in neuronal populations sensitive to a stimulus is affected after repetition of that stimulus (Grill-Spector, Henson, & Martin, 2006; Krekelberg, Boynton, & Wezel, 2006). fMRI adaptation can also be used to identify areas sensitive to particular stimulus attributes, by manipulating repetition of different attributes independently. Importantly, in such a paradigm multiple simultaneously occurring stimulus properties can be spatially segregated.

In Menenti et al. (2011) we applied this logic to investigate the global overlap between speaking and listening in three important components of language: semantic, lexical and syntactic processing. By comparing fMRI adaptation effects for semantic, lexical, and syntactic repetition in speaking and listening, we found that for all three components, the neuronal infrastructure was largely shared between speaking and listening. Bilateral posterior middle temporal gyri were involved in sentence-level semantic processing. Left posterior and anterior middle temporal gyrus, and left inferior and middle frontal gyrus, and the homologous areas on the right, were involved in lexical processing. Left posterior middle temporal gyrus and left inferior frontal gyrus were involved in syntactic processing.

In that study, however, we left unanswered the question of exactly how all the areas we found involved conspire in producing, or understanding, an utterance. For comprehension, this issue has been extensively addressed in the literature (Bookheimer, 2002; Friederici, Ruschemeyer, Hahne, & Fiebach, 2003; Hagoort, 2005; Martin, 2003). For production, however, our study provided the first useful set of data to look at the interplay of these components in more detail.

In this paper, we investigate the neuronal infrastructure for different steps in speech production. We look at overlap and segregation of semantic, lexical and syntactic processing in the brain, to investigate how the whole process of speaking, from thought to spoken words, may be performed through the cooperation of a network of brain areas.

In our study, 20 Dutch participants described pictures by producing short sentences, while lying in the MRI-scanner. Syntactic, semantic or lexical aspects of the spoken sentences could be either repeated or novel between two subsequent sentences (Fig. 1). Syntactic repetition is manipulated independently of both lexical and semantic repetition, enabling us to find areas uniquely sensitive to syntax. Lexical and sentence-level semantic repetition cannot be manipulated orthogonally to each other, as it is hard to communicate the same message without, at least partly, using the same words. Our design, however, allows us to manipulate repetition of word meaning and thematic role structure respectively, while keeping the other constant. This allows us to distinguish areas sensitive to word meaning from areas involved in sentence-level semantic processing. To counter MR-artefacts due to speaking, we used an fMRI-sequence that increases sensitivity and reduces motion artefacts (Buur, Poser, & Norris, 2009; Poser, Versluis, Hoogduin, & Norris, 2006).

2. Methods

For experimental methods, also see Menenti et al. (2011).

2.1. Subjects

Twenty-four (seven male) healthy right-handed (as assessed through an adapted version of the Edinburgh Handedness Inventory (Oldfield, 1971)) Dutch native speakers with normal or corrected to normal vision (mean age 22 years, range 20–28) participated in the experiment. Four subjects were excluded from analysis (one male) due to technical problems with the data.

2.2. Stimuli

Our target stimuli were photographs that depicted 36 transitive events such as *kiss*, *help*, *strangle* with the agent (doer) and patient (undergoer) of this action. These pictures could be described with transitive sentences, which were spoken by the participants. The pictures were displayed with one actor colored in red and the other in green, to cue the participants in our *stoplight paradigm* (see below). Besides the target stimuli, we also had filler pictures that depicted either intransitive or locative events.

2.3. Design

The design is illustrated in Fig. 1. There were three factors (syntax, semantics, words), with two levels (repeated/novel) each. The syntactic structure of subsequent sentences could either be the same (e.g., active–active) or different (e.g., active–passive). Separate repetition of meaning (semantic repetition) and individual words (lexical repetition) allowed us to distinguish areas that are sensitive to sentence meaning, and those that are sensitive to repetition of words but not to repetition of sentence meaning. We could not do so by using orthogonal factors, since it is not possible to repeat sentence meaning without, at least partly, repeating words. This effectively led to a $2 \times 2 \times 2$ design (semantics repeated/novel, words repeated/novel, syntax repeated/novel) with two empty cells (semantics repeated – words novel, for both novel and repeated syntax). We analyzed the effect of semantic repetition only in those conditions where the lexical content was repeated (thereby keeping the factor words constant), and the effect of word repetition only for those conditions where semantic structure was novel (therefore, keeping the factor meaning constant). Syntactic repetition was orthogonal to these factors.