Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Special issue: Research report

How do speakers coordinate? Evidence for prediction in a joint word-replacement task

Chiara Gambi [a,*], Uschi Cop [b] and Martin J. Pickering [a]

[a] Department of Psychology, University of Edinburgh, Edinburgh, UK
[b] Department of Experimental Psychology, Ghent University, Ghent, Belgium

ARTICLE INFO

Article history:
Received 17 April 2014
Reviewed 4 June 2014
Revised 4 July 2014
Accepted 10 September 2014
Published online 28 September 2014

Keywords:
Coordination
Joint task
Prediction
Forward model
Error repair

ABSTRACT

We investigated whether speakers represent their partners' task in a joint naming paradigm. Two participants took turns in naming pictures; occasionally the (initial) picture was replaced by a different picture (target), signaling that they had to stop naming the initial picture. When the same participant had to name the target picture, he or she completed the name of the initial picture more often than when neither participant had to name the target picture. Crucially, when the other participant had to name the target picture, the first participant also completed the name of the initial picture more often than when neither participant named the target picture. However, the tendency to complete the initial name was weaker when the other participant had to name the target than when the same participant went on to name the target. We argue that speakers predict that their partner is about to respond using some, but not all, of the mechanisms they use when they prepare to speak.

1. Introduction

There is substantial evidence that observers predict actions (e.g., Kilner, Vargaa, Duval, Blakemore, & Sirigu, 2004; Ramnani & Miall, 2004; Flanagan & Johansson, 2003; Graf et al., 2007; see Wilson & Knoblich, 2005 for a review). For example, the readiness potential, which indexes the preparation of motor responses, is present from about 500 msec prior the observation of a predictable hand action (Kilner et al., 2004). Similarly, comprehenders often predict language (e.g., Altmann & Kamide, 1999; Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; see Huettig, Rommers, & Meyer, 2011; Kutas, DeLong, & Smith, 2011; Pickering & Garrod, 2007; Van Petten & Luka, 2012 for reviews and discussion). For example, readers experience difficulty (i.e., enhanced N400) when the form of the indefinite article in English is not consistent with the initial phoneme of a highly expected noun (e.g., "an" when the expected noun begins with a consonant; DeLong, Urbach, & Kutas, 2005), indicating that phonological features of an upcoming word can be predicted.

But how do comprehenders compute such predictions? Researchers have proposed different mechanisms (Kutas, et al., 2011; Levy, 2008; Pickering & Garrod, 2007, 2013). In

* Corresponding author. Psycholinguistics Group, Department of Computational Linguistics and Phonetics, Building C 7.4, Campus Saarbruecken, Saarland University, Saarbruecken 66123, Germany.
E-mail addresses: g.chiara86@gmail.com, gambi@coli.uni-saarland.de (C. Gambi), uschi.cop@ugent.be (U. Cop), martin.pickering@ed.ac.uk (M.J. Pickering).
http://dx.doi.org/10.1016/j.cortex.2014.09.009
0010-9452/

this paper, our aim is to answer one general question about the nature of such mechanisms, that is: To what extent are the mechanisms used for prediction related to the mechanisms used when preparing to speak? In other words, are the process of preparing to speak and the process of predicting whether another person is about to speak related to one another? If so, one would expect predictions to affect language production on-line. More precisely, if the same mechanism is implicated concurrently in speech preparation and in predicting that another person is about to speak, then we would expect the latter process to affect the former.

There is some evidence that production processes might be involved in prediction during language comprehension. Federmeier, Kutas, and Schul (2010) reported that a late prefrontal positivity induced by plausible but unexpected nouns (which is thought to index error correction and/or prediction updating; Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007) is greatly reduced in older compared to younger adults. Importantly, the magnitude of this component in the older group correlated with production measures of verbal fluency (see also DeLong, Groppe, Urbach, & Kutas, 2012). Similarly, Mani and Huettig (2012) found that 2-year-olds with larger production (but not comprehension) vocabularies were more likely to predict upcoming referents (by looking at corresponding pictures) than their peers with smaller production vocabularies. These studies suggest that the ability or tendency to predict during language comprehension is correlated with language production abilities both in older adults and in children.

Pickering and Garrod (2013) proposed that prediction during language comprehension is subserved by the same mechanism that subserves feedforward control during language production, namely forward models (e.g., Wolpert, 1997). In their proposal, forward models map from production commands (communicative intentions) to the (production and comprehension) representations that will be retrieved as a consequence of executing those production commands. During language production, forward-model predictions are used for self-monitoring and learning. During comprehension, they are used in other-monitoring, and crucially to speed up and enhance understanding of the speaker's utterances (see Pickering & Garrod, 2014).

Recent MEG evidence suggests that covert language production (imagining to articulate or covert rehearsal in working memory) can selectively enhance early auditory responses to syllables (Tian & Poeppel, 2013; Ylinen et al., 2014). In addition, motor activation occurs during speech perception, particularly during adverse conditions (D'Ausilio, Bufalari, Salmas, & Fadiga, 2012). Finally, activation in the right cerebellum correlates with adaptation to distorted speech in a perceptual task (Guediche, Holt, Laurent, Lim, & Fiez, 2014), while rTMS of the right cerebellum delays predictive eye-movements to upcoming linguistic referents (Lesage, Morgan, Olson, Meyer, & Miall, 2012). Importantly, the cerebellum has been implicated in the computation of forward models by several authors (e.g., Ito, 2008), and there is some evidence that the computation of motor-to-auditory mappings might be atypical in patients with cerebellar lesions (Knolle, Schröger, & Kotz, 2013).

In sum, there is converging evidence for the implication of production mechanisms in prediction of one's own and others' utterances. Specifically, the evidence reviewed above suggests that prediction could involve some form of internal simulation of a production process, and that it might be remarkably specific. In other words, comprehenders might simulate, using language production mechanisms, details of the linguistic content of another's utterance, for example associated with meaning (e.g., such as whether an upcoming referent is likely to be an edible object) or sound (e.g., whether an upcoming noun is likely to start with a consonant, or whether an upcoming vowel is likely to involve formant frequencies within a certain range).

However, at present, clear causal evidence for the implication of language production processes in content-specific prediction is limited to phonetics. Neurophysiological studies of syllable or pseudoword perception show that an articulation-related mechanism (i.e., activation of speech motor programs) is responsible for the effects of overt and covert language production on neural responses in auditory areas. But we do not yet know whether the same would hold for words and other meaningful units.

In addition, production and prediction could share general-purpose mechanisms (e.g., heightened attention; preparedness to respond), rather than language-specific mechanisms (i.e., processes involved in formulating utterances). While general-purpose mechanisms would not be able to support prediction of specific linguistic content (i.e., what somebody is about to say), they could in principle support prediction of whether another is about to speak (or, indeed, act in some other way). Such mechanisms could, for example, help speakers to predict whether another conversational participant is about to take the floor (Wenke et al., 2011), either by producing a linguistic utterance or by producing a non-verbal utterance (e.g., pointing gesture; Clark, 1996).

In this study, we tested the hypothesis that speakers predict whether another person is about to speak using mechanisms that are also implicated when they prepare to speak themselves, and investigated whether beliefs about another person's upcoming task can affect the way a speaker produces his or her own utterance. To this aim, we devised a joint language production task that requires participants to take turns in speaking. Joint tasks have been used to study similar issues in the domain of action, as we briefly discuss below before returning to language.

1.1. *Joint tasks in the action domain*

In joint task paradigms (Knoblich, Butterfill, & Sebanz, 2011; Sebanz, Bekkering, & Knoblich, 2006; Sebanz & Knoblich, 2009), participants are tested in pairs and are assigned complementary tasks (i.e., they each perform half of the task that would be performed by a single participant in solo task paradigms; see below). For example, in one study (Knoblich & Jordan, 2003), pairs of participants attempted to keep a circle aligned with a moving dot on a computer screen. In each pair, one participant could accelerate the tracker only to the right, while the other could accelerate it only to the left. Performance in the joint task is usually compared to performance in a solo version of the same task. In this study, in the solo version an individual participant could control the tracker's velocity in both directions (using two hands).