Cognition 133 (2014) 395–407

Contents lists available at ScienceDirect

Cognition

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

On predicting others' words: Electrophysiological evidence of prediction in speech production

Cristina Baus [a,b,*], Natalie Sebanz [c], Vania de la Fuente [a], Francesca Martina Branzi [a], Clara Martin [d,e], Albert Costa [a,f]

[a] Center of Brain and Cognition, CBC, Universitat Pompeu Fabra, Barcelona, Spain
[b] Laboratoire de Psychology Cognitive, CNRS – Universite Aix-Marseille, Marseille, France
[c] Department of Cognitive Science, Central European University, Budapest, Hungary
[d] Basque Center on Cognition, Brain and Language, BCBL, Donostia, Spain
[e] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
[f] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

ARTICLE INFO

Article history:
Received 5 March 2013
Revised 5 June 2014
Accepted 21 July 2014

Keywords:
Joint action
Speech production
Lexical frequency

ABSTRACT

The present study investigated whether lexical processes that occur when we name objects can also be observed when an interaction partner is naming objects. We compared the behavioral and electrophysiological responses of participants performing a conditional go/no-go picture naming task in two different conditions: individually and jointly with a confederate participant. To obtain an index of lexical processing, we manipulated lexical frequency, so that half of the pictures had corresponding names of high-frequency and the remaining half had names of low-frequency. Color cues determined whether participants should respond, whether their task-partner should respond, or whether nobody should respond. Behavioral and ERP results showed that participants engaged in lexical processing when it was their turn to respond. Crucially, ERP results on no-go trials revealed that participants also engaged in lexical processing when it was their partner's turn to act. In addition, ERP results showed increased response inhibition selectively when it was the partner's turn to act. These findings provide evidence for the claim that listeners generate predictions about speakers' utterances by relying on their own action production system.

1. Introduction

Many joint actions require that we anticipate others' actions: Think of playing a piano duet, dancing a tango or walking through a narrow doorframe together. It is quite easy to imagine the consequences of not predicting in advance whether our partner will take the first turn crossing the doorframe or whether he/she will leave the first turn to us. The same is true for a conversation, which constitutes a paradigm case of joint action (Clark, 1996; Garrod & Pickering, 2004). When having a conversation, predicting others' verbal actions and integrating them in our own action plan is key. In the present article we explore the involvement of the production system in predicting another's verbal actions.

It has been suggested that predicting others' actions involves processes that are also engaged in the planning and performance of one's own actions (e.g., Knoblich, Butterfill, & Sebanz, 2011). There is considerable evidence for the engagement of motor representations not only during action perception (e.g., Rizzolatti & Craighero, 2004)

* Corresponding author. Address: Department de Tecnología, Universitat Pompeu Fabra, C/Tànger, 122-140, 08018 Barcelona, Spain. Tel.: +34 93 542 13 82; fax: +34 93 542 25 17.

E-mail address: baus.cristina@gmail.com (C. Baus).

http://dx.doi.org/10.1016/j.cognition.2014.07.006

but also during anticipation of others' actions (Aglioti, Cesari, Romani, & Urgesi, 2008; Kourtis, Sebanz, & Knoblich, 2013; Ramnani & Miall, 2003; van Schie, Mars, Coles, & Bekkering, 2004). This evidence supports the assumption that interaction partners predict each other's actions through motor simulation (Wilson & Knoblich, 2005). Forward models in the motor system allowing one to predict one's own actions may also enable the prediction of others' actions at multiple levels (Wolpert, Doya, & Kawato, 2003; see also, Brown & Brüne, 2012, for a review).

This interweaving of action and action perception has recently been spelled out for the role of language production in conversational contexts. According to Pickering and Garrod (2007) (see also, Gambi & Pickering, 2011; Pickering & Garrod, 2013, for reviews) the language production system generates forward models (imitative plans) at specific levels of representation, including semantics, syntax, and phonology, to predict utterances during language comprehension. However, apart from studies revealing the involvement of motor processes during speech perception (e.g., Fadiga, Craighero, Buccino, & Rizzolatti, 2002; *motor theories of speech perception*, see Hickok, 2012), little is known about the exact role that our production system plays in predicting others' utterances, especially at more abstract levels of representation (e.g., lexical) where articulation is not present. The main objective of the present study was to investigate whether lexical processes in speech production are involved during the anticipation of a task-partner's utterance. To achieve this aim, a task sharing paradigm (see below) was adapted to a picture naming task (hereafter *joint picture naming task*) in which pictures had to be named either by a participant, by a task-partner, or by no one (depending on the color in which the pictures were presented).

Task sharing paradigms have been used to study a particular process of joint action, namely action planning (Knoblich et al., 2010). Very briefly, this experimental approach consists in two individuals performing independent tasks in a shared setting. Importantly, since participants are not explicitly required to coordinate their actions, the task sharing paradigm provides a conservative estimate of the extent to which people engage in planning not only their own actions, but also their task-partner's actions.

The main observation from task sharing studies has been that two individuals performing one part of a task each show a similar pattern of performance as one individual performing both parts on her own (e.g., Sebanz, Knoblich, & Prinz, 2003; Welsh, 2009). This was first demonstrated using a spatial compatibility task (Simon, 1990) where participants are instructed to respond to stimulus color (e.g., left key press for red and right key press for green stimuli) and to ignore stimulus location. When participants perform this task alone, they show faster responses when the irrelevant stimulus location and the spatial location of the response to be given to stimulus color overlap. Sebanz et al. (2003) developed a social version of the compatibility task and compared participants' performance in two conditions: an individual go/no-go condition in which the participant was alone in the room and was instructed to respond only to one of the colors (e.g. respond to red) and to do nothing for the other color (e.g. do not respond to green) and a joint go/no-go condition in which participants performed the task with a partner. Importantly, the task for the participant was the same as in the individual go/no-go condition (e.g. respond only to red). The only difference was that the partner was instructed to perform the complementary task (e.g. respond to green). Sebanz et al. (2003) showed a spatial compatibility effect in the joint condition (similar to the standard individual condition in which the participant was instructed to perform both tasks), but not when the participant was performing the task alone. This has been taken to indicate that our own actions and others' actions are planned in a functionally similar manner (Welsh, 2009).

Electrophysiological studies have also employed the task sharing paradigm to investigate what happens during no-go trials that do not require a response from the participant, but from his or her task-partner. The relevant comparison here is between these no-go trials the task-partner needs to respond to and a second set of no-go trials that neither the participant nor the task-partner needs to respond to. Larger amplitudes of the so-called No-Go P300 (actually peaking around 450–550 ms after stimulus onset; Sebanz, Knoblich, Prinz, & Wascher, 2006; Tsai, Kuo, Hung, & Tzeng, 2008; Tsai, Kuo, Jing, Hung, & Tzeng, 2006) have been reported for no-go trials that require the task-partner to respond compared to those that nobody responds to. This No-Go P300 modulation has been taken as an index that the other's action is planned, and that, consequently, inhibitory action control processes are required to ensure that participants do not act when it is the other's turn.

Although several different versions of the task sharing paradigm have been developed and yielded a rich set of behavioral and electrophysiological findings (for an overview, see Obhi & Sebanz, 2011; Wenke et al., 2011), it still remains unclear which aspects of the other's task are included in our own planning. The crucial question is to what extent people mentally perform the other's task when it is not their own turn, but their task-partner's turn to act. According to the *actor co-representation* account (Dolk et al., 2011; Philipp & Prinz, 2010; Vlainic, Liepelt, Colzato, Prinz, & Hommel, 2010; Wenke et al., 2011), task-partners form a representation that specifies which events they are responsible for and which events require their partner to act (e.g., red: me; green: you). The *task co-representation* account (Atmaca, Sebanz, & Knoblich, 2011; Sebanz, Knoblich, & Prinz, 2005) claims that representations of another's task specify not only *when* the other needs to act, but also *what* she needs to be doing (e.g., green: task-partner needs to press right key). Despite several attempts, previous studies have largely failed to find conclusive evidence for task co-representation since the joint compatibility effect described above can be explained as a result of representing when it is the co-actor's turn (Philipp & Prinz, 2010; Wenke et al., 2011) or even just being sensitive to her spatial location (Dittrich, Dolk, Rothe-Wulf, Klauer, & Prinz, 2013; Dolk et al., 2011).

The fact that unlike other forms of action, language is inherently social and has been very well characterized in terms of levels of processing (e.g., lexical) and their electrophysiological correlates (e.g., lexical frequency), makes it highly suitable to explore prediction processes. In particular, it provides a new route for addressing the question of whether task-partners predict the specific actions to be performed by their partners, as predicted by the task co-representation account. Will a partner/listener who represents a speaker's task engage in similar lexical processes as the speaker? As a first step towards addressing this question, we investigated whether task-partners performing a joint naming task would engage in lexical processes when it was their partner's turn. If so, we should see similar markers of lexical processing in the speaker and the listener. Specifically, we explored whether lexical frequency effects (as a marker of lexical processing) typically observed in object naming tasks when a person is asked to perform the task alone are also present when the same task is shared by two individuals.

In speech production research, there is ample evidence showing that the speed and accuracy with which items are retrieved from the lexicon is influenced by their lexical frequency. For instance, picture naming studies have repeatedly shown that pictures whose corresponding names are of high lexical frequency are named faster and more accurately than those pictures whose corresponding names are of low lexical frequency (e.g., Almeida, Knobel, Finkbeiner, & Caramazza, 2007; Caramazza, Bi, Costa, & Miozzo, 2004; Caramazza, Costa, Miozzo, & Bi, 2001; Dell, 1990; Jescheniak & Levelt, 1994; Kittredge, Dell, Verkuilen, & Schwartz, 2008; Navarrete, Basagni, Alario, & Costa, 2006; Wingfield, 1968). At the electrophysiological level, frequency effects typically manifest around 200 ms after the picture onset presentations. Low-frequency words elicit larger amplitudes than high-frequency ones, especially at posterior sites (Sahin, Pinker, Cash, Schomer, & Halgren, 2009; Strijkers, Baus, Runnqvist, Fitzpatrick, & Costa, 2013; Strijkers, Costa, & Thierry, 2010; Strijkers, Holcomb, & Costa, 2011). Modulations of the P200 ERP component for word frequency as well as for other lexical phenomena (e.g., cognates, semantic competitors; e.g., Costa, Strijkers, Martin, & Thierry, 2009) have been taken as an index of (or sensitivity to) lexicalization processes. Larger amplitudes are elicited by those words overall less active in the lexicon (e.g., low-frequency words, non-cognates) and therefore more difficult to retrieve.

The reported frequency effects offer us clear predictions of what we might expect in the present study for those trials (go trials) in which the participant is required to name the pictures. But will a frequency-driven ERP modulation be present when the participant does not need to name a picture, but her task partner does (i.e., no-go trials)? The task co-representation account predicts an effect of lexical frequency because it assumes that participants have formed a representation of the specific task to be performed by the other, which may trigger predictions of the other's utterance.

As the extent to which words become activated during no-go trials cannot be known in advance, in the present study we adopted the same strategy as in previous ERP studies on task sharing (e.g., Sebanz et al., 2006). In addition to no-go trials assigned to the task-partner, a second set of no-go trials was included that required neither the response of the participant nor of the task-partner. As both types of trials are of the same nature (no-go trials), differences between them regarding the frequency effect can only be the result of participants predicting the task partner's upcoming words.

In the following, we describe the specific details of the present study and the expected results regarding the presence of frequency effects in the different conditions.

1.1. The joint picture naming task

In this study, participants were asked to perform a conditional picture naming task. They were presented with pictures in three different colors and they were asked to name only pictures in one of the three colors (go trials; e.g., red). Hence there were two types of no-go trials (e.g., pictures in blue and pictures in black).

Crucially, the experimental session was split into two parts. In the individual condition, participants performed the task alone in the experimental room. In the joint condition, participants performed the task together with a task-partner (i.e., confederate) who was sitting alongside the participant in the same room. In both conditions participants were asked to do the same: name the pictures in one color (e.g., red) and do nothing when the pictures appeared in the other two colors. Importantly, in the joint condition, the confederate participant was asked to name the pictures that appeared in one of the colors (e.g., blue), and to do nothing when the pictures appeared in a different color (e.g., in red or in black). Hence in the joint condition, from the participant's perspective, there were three different types of trials: (a) trials in which the participant was supposed to name the pictures (*self-go trials*), (b) trials in which the participant was not supposed to name the picture but the confederate participant was (*other-go trials*), and (c) trials in which none of the participants named the pictures (*joint no-go trials*).

Of particular interest for our study were the two types of trials in which the participant was not required to respond. We compared the ERP response on no-go trials in the joint condition where none of the two participants responded (*joint no-go trials*) with the ERP response on those no-go trials where the task-partner performed the naming task (*other-go trials*). Both of these no-go trials are identical from the participant's perspective, the crucial difference being whether the confederate names the picture or not. By comparing the participants' brain activity for these two types of no-go trials we can assess the extent to which the participants covertly perform the action to be carried by their task-partner. Firstly, we expect to replicate earlier findings revealing larger amplitudes (No-go P300, around 450–550 ms post stimulus onset; e.g., Sebanz et al., 2006) for those trials that require the task partner's action, compared to no-go trials that do not require anyone's response. Secondly, if participants predict the upcoming word that their partner will produce through activation of their own production system (Pickering & Garrod, 2007), then differences in the ERP components associated with frequency for the two types of no-go trials should also be observed. Thus, the crucial issue is whether frequency effects will appear in the participant for those trials named by the confederate.

Given the novelty of this approach, we cannot be sure whether the latency of the frequency effect on no-go trials will be the same as reported in object naming studies. In particular, we do not know to what extent withholding a response, which is cognitively more demanding than responding (e.g., Bokura, Yamaguchi, & Kobayashi, 2001), might alter the time course of lexical access during word production (Strijkers et al., 2011). Little is known about how to-be-ignored stimuli are processed in the brain and what stimulus properties might affect the depth of linguistic processing (e.g., Bles & Jansma, 2008). Accordingly, the comparison between the two types of no-go trials becomes especially relevant since the task of the participant is exactly the same for both types of trials, namely to do nothing.

In sum, in the present study we aimed at exploring task co-representation in the context of speech production. Specifically, we investigated whether lexical processes that occur when a participant names an object can also be observed when predicting a task partner's utterances.

2. Method

2.1. Participants

Thirty-six Spanish native speakers (mean age = 22.3 (SD = 3); 19 women) took part in the experiment. All of them were right-handed and had normal or corrected to normal vision. Four Spanish female confederates acted as task-partners. Participants were informed that they would perform the corresponding part of the experiment with another person just before the confederate was asked to enter the room. EEG was only registered for the participant.

2.2. Materials

Two sets of 150 pictures belonging to different semantic categories were selected from different picture databases (e.g., Bates, D'Amico, Jacobsen, Székely, et al., 2003; Snodgrass & Vanderwart, 1980). Each set was randomly assigned to the individual or the joint condition for a given participant. The frequency value of each picture name was extracted from a Spanish word database with 31,491 lexical entries (BuscaPalabras; Davis & Perea, 2005). Within each set of pictures, half had high frequency names (e.g., sun, table: set 1: mean = 47 occurrences per million, sd = 70; set 2: mean = 43 per million, sd = 56) and the other low frequency names (e.g., zebra, pineapple; set 1: mean = 3.4 per million, sd = 2; set 2: mean = 3.5 per million, sd = 2).

In each condition, 50 pictures were assigned as go trials and 100 pictures as no-go trials (in the joint condition, 50 of these were 'other-go' trials). Half of the go trials and half of the no-go trials contained high-frequency pictures and half low-frequency pictures. Pictures appeared in red, black, and blue to indicate whose turn it was. The assignment of colors to participants was counterbalanced across participants.

2.3. Procedure

The experiment was divided into two blocks, corresponding to the individual and joint condition. The order of these blocks was counterbalanced across participants, so that half of the participants started the experiment with the individual condition and the other with the joint condition.

In the individual condition, participants were alone in the room and sat in front of the computer screen. They were asked to name the pictures appearing in a given color and to do nothing when the pictures appeared in a different color. At the beginning of the block, a colored rectangle (in red, black or blue) was presented on the screen indicating to the participant her assigned color. The assigned color for a given participant was maintained throughout the two experimental conditions, individual and joint. Naming latencies and error rates were measured for the go trials. Participant's brain activity was registered continuously during the experiment, hence including go and no-go trials.

In the joint condition, the participant and the confederate sat alongside each other in front of the computer screen. Both participants were given instructions together. The experimenter told them that their task was to name only those pictures appearing in their assigned color and to do nothing for the rest of the pictures. Participant and confederate were informed that only one picture would be presented at a time, so no picture had to be named by both of them at the same time. The corresponding colors for the confederate and the participant were assigned at the beginning of the experiment by means of two colored rectangles presented on the right and left hand side of the screen. After the joint condition was finished, the confederate was asked to leave the room.

Trial structure was as follows: a fixation point (*) was presented in the middle of the screen for 1500 ms followed by the picture presentation. Pictures on no-go trials in the individual condition and the joint no-go trials in the joint condition were presented on the screen for a random duration between 800 and 1200 ms, to roughly match the presentation time for pictures on go trials. Those pictures that had to be named either by the participant (*self-go*) or by the confederate (*other-go trials*) were presented until a response was given or for 3000 ms maximum. Naming latencies were measured from the onset of the picture presentation.

Once the experiment had finished, participants were presented with the pictures previously assigned to the confederate. In order to check whether the participant would have named the pictures as the confederate did, they were instructed to name these pictures, regardless of the answers given previously by the confederate. Only those responses that matched between the participant and the confederate (both providing the same name for a given object) were included in the ERP analysis.

2.3.1. ERP recording

EEG was continuously registered and linked-nose reference from 31 scalp Ag/Cl passive electrodes. Two external electrodes were placed at right and left mastoids. Eye movements were monitored by two external electrodes placed horizontally (outer canthus) and vertically (below) to the right eye. The impedance of the electrodes was kept below 5 kΩ (10 kΩ for the ocular electrodes). EEG signal was digitalized online with a 500 Hz sampling rate and a band pass filter of 0.1–125 Hz. EEG data was filtered offline to 0.03 Hz high-pass filter and 20 Hz low-pass filter and vertical and horizontal ocular artefacts were corrected by a correction algorithm (Gratton, Coles, & Donchin, 1983). Afterwards, data was segmented into 750 ms epochs (−200 to 550 ms). Before averaging, segments with incorrect responses, containing artefacts (brain activity above or below 100 μV or a change in amplitude between adjacent segments of more than 200 μV) or eye blinks were excluded. Epochs were then averaged in reference to −100 ms pre-stimulus baseline. As the go condition was more affected by the motor artefacts generated by speech articulation, we only included those participants in the analysis for whom more than 50% of the segments in the self go condition could be retained (average segments per condition: 19.8). This led to the exclusion of four participants from the analysis.

2.4. Data analysis

We analyzed participants' behavioral (for go trials) and ERP responses (for go and no-go trials).

2.4.1. Behavioral analysis

For go trials, no responses and hesitations (mmm, uh) were considered as errors and excluded from the naming latency analysis. Moreover, verbal responses different from those we had intended to elicit were excluded from the analysis because even though they were plausible words (e.g., naming *bird* for the picture of a *parrot*), they differed from the target word with respect to their frequency value (e.g., parrot in Spanish has a frequency value of 5, falling in the range of low-frequency words, while bird has a frequency of 20, falling in the range of high-frequency words). After excluding participants with more than 25% of the pictures incorrectly named (two participants) or with many artefacts in the ERPs (four participants), the final sample included 30 participants.

Naming latencies and error rates were analyzed by fitting Generalized Linear Mixed Effects models with the lme4 library in R (Bates, Maechler, & Bolker, 2011; see also Baayen, 2008; R Development Core Team, 2013). Naming latencies were log-transformed (for a better fit of the model) and latencies three standard deviations below or above the participant's mean were excluded from the analysis. Errors were analyzed by fitting a logistic model, more suitable to analyze binary data (Jaeger, 2008). The analyses included frequency (high-frequency vs. low-frequency) and condition (individual vs. joint) as fixed factors and participants and items as random factors. Both lexical frequency and condition were contrast coded. Mean log-naming latencies and error rates for high frequency words at the individual condition were taken as the intercept (baseline condition) against which the other conditions were compared. The *t*-values for the coefficient and probability (p_{mc}) values were based on 10,000 Markov Chain Monte Carlo (MCMC) samples (Baayen, Davidson, & Bates, 2008).

2.4.2. ERP analysis

Based on previous evidence of the electrophysiological correlates of lexical processing in speech production, we focussed on two time-windows: the 160–240 ms time-window (corresponding to the P200) and the 320–420 ms time-window (corresponding to the P300). The maximal peak across electrodes in each condition fell within the time range of the P200 (Average peak for Self go trials: 201 ms, SD = 32; No-go trials: 211, SD = 29). In contrast, in the time range of the P300 from 300 to 550, self-go and no-go trials differed in their maximal peaks (Self go trials: 370 ms, SD = 10; No-go trials: 448 ms, SD = 25 ms). Thus, a further time-window from 420 to 550 ms was also analyzed (late portion of the P300), primarily to test predictions concerning the No-go P300.

Go and no-go trials were analyzed separately. Analysis of the go trials served as a baseline to test the chronometry of lexical processing when speaking. To do so, a $2 \times 2 \times 3 \times 3$ ANOVA was conducted with Condition (Individual vs. Joint), word frequency (high vs. low frequency words), Region (Anterior, Central and Posterior) and Laterality (Left, Central, Right) (AnteriorLeft: F7, F3, FC5; AnteriorCentral: Fz, FC1, FC2; AnteriorRight: F8, F4, FC6; CentralLeft: T3, C3, CP5; CentralCentral: Cz, CP1, CP2; CentralRight: T4, C4, CP6; PosteriorLeft: P7, P3, O1; PosteriorCentral: Pz, PO1, PO2; PosteriorRight: P8, P4, O2).

Regarding the no-go trials, three different analyses were conducted: (1) Analyses comparing frequency effects in the two types of no-go trials in the individual condition. It should be made clear, that from the perspective of the participant, these two types of trials were the same (no-go trials). Thus, we acknowledge that in the individual condition type of trial can be seen as a one-level factor rather than a factor with two levels. However, this analysis is important to confirm that a priori there were no differences between the two types of no-go trials in the individual condition. This then implies that any difference between these two types of trials in the joint condition can only be attributed to the fact that the confederate was naming some of the pictures. An ANOVA with 2 (Type of no-go trial: individual no-go vs. other no-go) $\times$ 2 (word frequency: high vs. low frequency words) $\times$ 3 (region) $\times$ 3 (laterality) was conducted. (2) Analyses of the no-go trials in the joint condition where we expected to find differences between the two types of no-go trials. A $2 \times 2 \times 3 \times 3$ ANOVA with Type of trial (other-go vs. joint no-go), Frequency (high vs. low frequency words), Region and Laterality was conducted. 3) No-go trials in both conditions were submitted to a $2 \times 2 \times 2 \times 3 \times 3$ ANOVA with Condition (individual vs. joint), Type of trial (other-go vs. joint no-go), Frequency (high vs. low frequency words), Region and Laterality, to further explore the frequency effect in the no-go trials. Greenhouse-Geisser correction (corrected degrees of freedom and probabilities are reported) and Bonferroni correction for multiple comparisons were applied when necessary.

Latency analyses were also conducted to explore the onset of the frequency effect in the self-go trials in the individual and joint conditions and the other-go trials (and its absence in the joint no-go trials). To do so, ERP elicited by pictures with high frequency names were compared to those elicited by pictures with low frequency names by running two-tailed paired *t*-tests at every sampling rate (2 ms). The onset of the frequency effect was taken as the first data point of a sequence of consecutive significant data points (below 0.05 level, FDR corrected; Benjamini & Hochberg, 1995).

3. Results

3.1. Behavioral results

The naming latency results revealed that low-frequency words were named slower than high-frequency words ($\beta = 0.11$; s.e. = 0.36, $t = 3.06$, $p_{mc} < .01$). Moreover, participants responded faster in the individual than in the joint condition ($\beta = 0.11$; s.e = 0.34, $t = 3.6$, $p_{mc} < .01$). Importantly, these two factors did not interact ($\beta = -0.02$; s.e = 0.02, $t = -1.3$, $p = .18$), showing that frequency effects were not modulated by the context in which the participant was naming the pictures (see Fig. 1).

The error rate analysis revealed no significant effect neither for frequency ($pz = .14$), nor for condition (pz = .6; as well as the interaction between them, $pz = .8$).

3.2. ERP results

3.2.1. Frequency effects on go trials

An ANOVA with 2 (Condition: Individual vs. Joint) × 2 (Word frequency: high vs. low frequency words) × 3 (region) × 3 (laterality) was conducted on the three time-windows of interest: 160–240 ms, 320–420 ms and 420–550 ms, corresponding to the P200, P300 and the late portion of the P300 respectively. Effects involving laterality were only considered in case of significant interactions with our variables of interest (condition, word frequency and region).

In the 160–240 time-window (P200), pictures whose corresponding names were of low frequency elicited a more positive waveform than those pictures whose

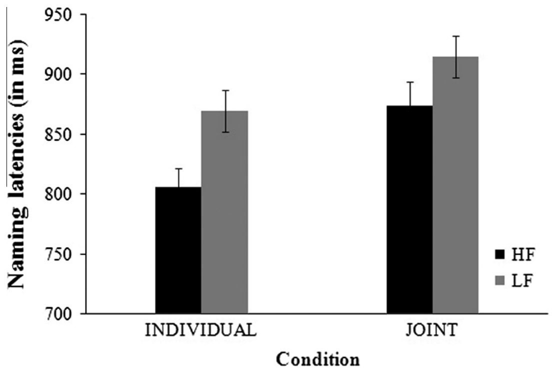

Fig. 1. Naming latencies for high and low frequency words for go trials in the individual and in the joint condition. The bars represent the standard errors.

corresponding names were of high frequency, as revealed by the main effect of lexical frequency ($F(1,29) = 4.6$, $p < .05$). The frequency effect was larger for those electrodes in the posterior region ($F(1.2,35.4) = 5.09$, $p < .05$, $\eta^2_p = .14$; post hoc paired-wise comparisons: Anterior: $p = .26$; Central: $p = .06$; Posterior: $p = .007$). No interaction with laterality was observed. Importantly, neither the main effect of condition nor its interaction with word frequency turned out to be significant (all $Fs < 1$). That is, in this time-window, frequency effects for go trials appear to be the same regardless of whether the participant was naming objects in the individual or in the joint condition (see Fig. 2).

Onset latency analysis revealed a difference between individual and joint conditions in the onset of the frequency effect. At the posterior region, the frequency effect started to be significant in the individual condition at 166 ms after the picture onset and remaining significant until 200 ms after the picture onset. In contrast, in the joint condition, the frequency effect started to be significant at 222 ms after the picture onset presentation and remained significant until 500 ms after the picture onset.

In the 320–420 ms time-window (P300), the frequency effect was only present over the posterior electrodes (Frequency * Region ($F(1.1,\ 33.1) = 8.4$, $p < .01$, $\eta^2_p = .19$; post hoc paired-wise comparison over the Posterior region: $F(1,29) = 5.8$, $p < .05$; Central region $p = .1$; Anterior region $p = .3$). The main effect of condition was not significant ($F < 1$). However, frequency effects (as indicated by low-frequency amplitudes being more positive than high-frequency ones) were significant in the joint condition but not in the individual condition (see Fig. 2). This is revealed by the significant interaction between frequency and condition ($F(1,29) = 4.1$, $p = .05$, $\eta^2_p = .14$), and the post hoc pairwise comparison of the frequency effects in the joint ($F(1,29) = 4.4$, $p < .05$) and in the individual condition ($F < 1$).

Analysis in the 420–550 ms time-window (late portion of the P300) revealed the same pattern of results reported in the previous time-window: the frequency effect was only significant in the joint condition ($p = .03$) but not in the individual one ($p = .17$) as revealed by the significant interaction between frequency and condition ($F(1,29) = 6.8$, $p < .05$, $\eta^2_p = .19$).

Overall, these analyses reveal several instructive findings. First, we were able to replicate the frequency effects for go trials observed previously in picture naming tasks. Moreover, the early effects of frequency (P200 time window) do not seem to be affected by whether or not the naming task is carried out alone or together with a confederate. However, the effects of frequency seem to be affected by the social context in which participants perform the task as revealed by the frequency effects observed only for the joint condition in the P300 time-window.

3.2.2. Frequency effects on no-go trials

In this analysis we explored the presence of lexical frequency effects for no-go trials.

Individual Condition: The first step was to assess whether there were any frequency effect for the no-go trials in the individual condition. To this aim, a first ANOVA

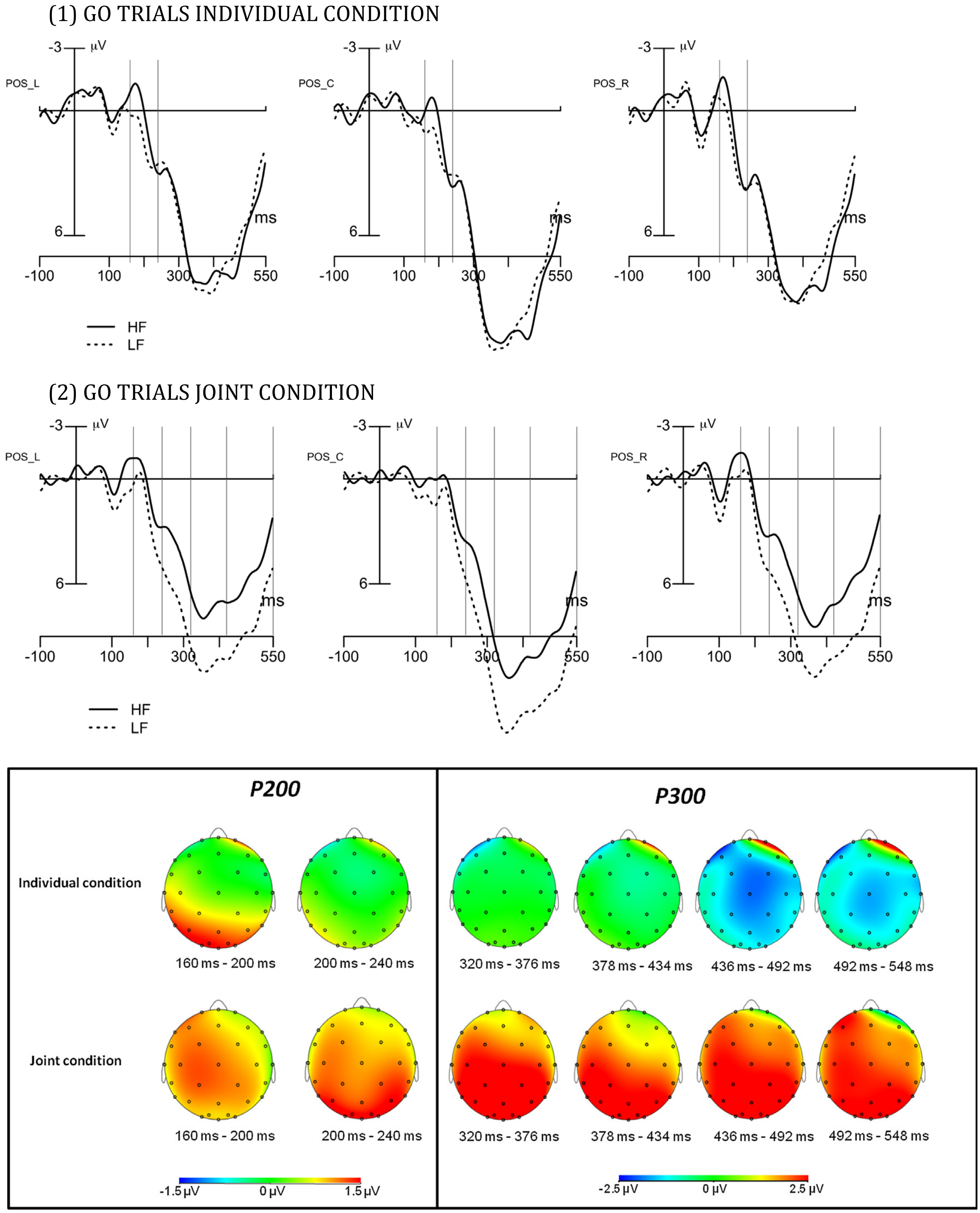

Fig. 2. EEG results in for the GO trials in the individual (1) and joint (2) conditions. The upper panels represent the grand averages for go trials in the individual (upper panel) and the joint condition (lower panel). Solid lines represent high-frequency words (HF) and dotted lines represent low-frequency words (LF). The three figures correspond to the linear derivation of the electrodes included the three posterior regions of interest: POS_L (posterior left: T5, P3 and O1), POS_C (posterior central: Pz, PO1 and PO2) and POS_C (posterior right: T6, P4 and O2). The gray vertical lines represent the time-windows in which the frequency effect was significant. The lower panel represents the topographical maps representing the frequency effect in the P200 and P300 time-windows (low frequency words minus high frequency ones). Positive differences (red colors) correspond to low frequency words being more positive than high-frequency ones. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

was conducted with 2 (Type of no-go trial: individual no-go vs. other no-go) × 2 (word frequency: high vs. low frequency words) × 3 (region) × 3 (laterality). In principle, if participants are following the instructions and ignoring the stimuli on the no-go trials, we should not observe any frequency effects (see Fig. 3). This is precisely what

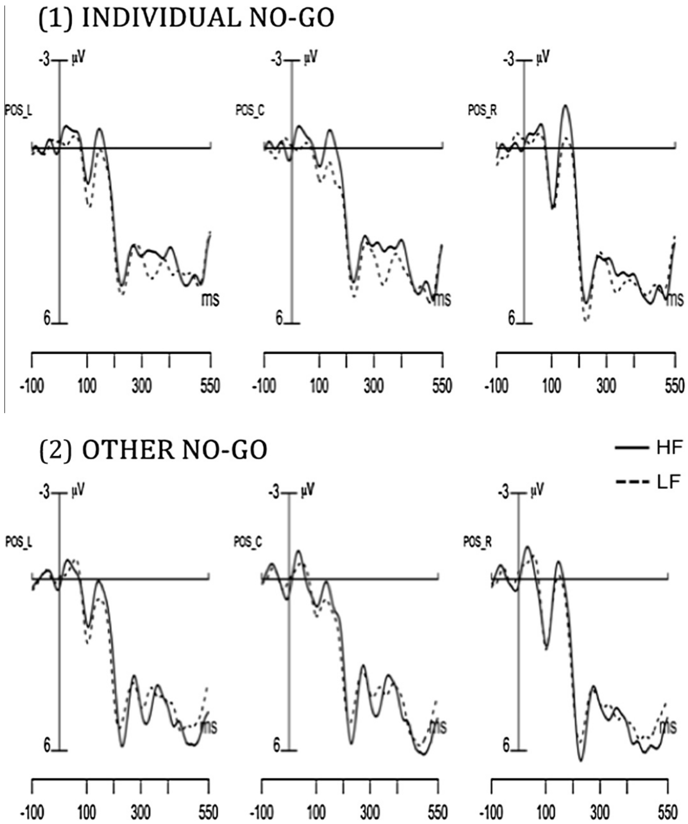

Fig. 3. EEG results for the two types of no-go trials in the Individual condition. The lines represent the grand averages for waveforms corresponding to the individual no-go (upper panel) and other no-go trials (lower panel). Solid lines represent high-frequency words (HF) and dotted lines represent low-frequency words (LF). The three figures correspond to the linear derivation of the electrodes included the three posterior regions of interest: POS_L (posterior left: T5, P3 and O1), POS_C (posterior central: Pz, PO1 and PO2) and POS_C (posterior right: T6, P4 and O2). The gray vertical lines represent the time-windows in which the frequency effect was significant.

we found, namely no frequency effects for the no-go trials in the individual condition, suggesting that indeed participants were not lexicalizing the pictures (at least not to the extent to show frequency effects) when they were not suppose to name them (P200 time-window: main effect of Frequency: $F(1,29) = 1.2$, $p = .2$, main effect of Type of no-go trial: $F(1,29) = 3.2$, $p = .08$ and Frequency * Type of no-go trial: $F < 1$) (P300 time-window: main effect of Frequency: $F < 1$, main effect of Type of no-go trial: $F < 1$ and Frequency * Type of no-go trial: $F < 1$) (late portion of the P300 (420–550 ms): main effect of Frequency: $F < 1$, main effect of Type of no-go trial: $F < 1$ and Frequency * Type of no-go trial: $F < 1$).

Joint Condition: More interesting is the comparison of the different no-go trials in the joint condition (joint no-go, where no one responds, vs. other-go, where the confederate but not the participant responds). An ANOVA with 2 (Type of no-go trial: joint no-go vs. other-go) × 2 (word frequency: high vs. low frequency words) × 3 (region) × 3 (laterality) was conducted in the same two time-windows.

In the 160–240 ms time-window (P200), we did not observe any modulation associated with frequency ($F < 1$). Neither the main effect of Type of no-go trial ($F(1,29) = 1.9$, $p = .17$, $\eta^2_p = .06$), nor the interaction with frequency were significant ($F < 1$). This indicates that brain activity for no-go trials in this time window is not modulated by whether or not the confederate is supposed to name the pictures (see Figs. 4 and 5).

In the 320–420 ms time-window (P300), there was a main effect of Frequency ($F(1,29) = 8.3$, $p < .01$, $\eta^2_p = .22$). The main effect of Type of no-go trial was not significant ($F < 1$). Crucially, however, the interaction between word frequency and type of trial was significant, revealing that type of trial exerted an influence on the presence of the word frequency effect ($F(1,29) = 4$, $p = .05$, $\eta^2_p = .12$). As predicted, the word frequency effect was only present for the other-go trials ($F(1,29) = 12.2$, $p < .01$, $\eta^2_p = .29$), but not for the joint no-go trials ($F(1,29) = 2$, $p = .16$, $\eta^2_p = .06$). This effect was significant over all regions but maximal over the posterior region (as indicated by the significant three-way interaction word frequency * type of no-go trial * region: $F(1.1,33.3) = 3.9$, $p = .05$) (see Figs. 4 and 5).[1]

[1] There were also differences in the timing associated with the frequency effects in the joint condition depending on whether the participant or the confederate were asked to name a picture. Onset latency analysis in the joint condition confirmed the early onset of the frequency effect observed when the participant had to name the object (222 ms after the picture onset presentation) relative to when the confederate was required to do it (342 ms after the picture onset presentation). The joint no-go condition did not reveal a significant frequency effect.

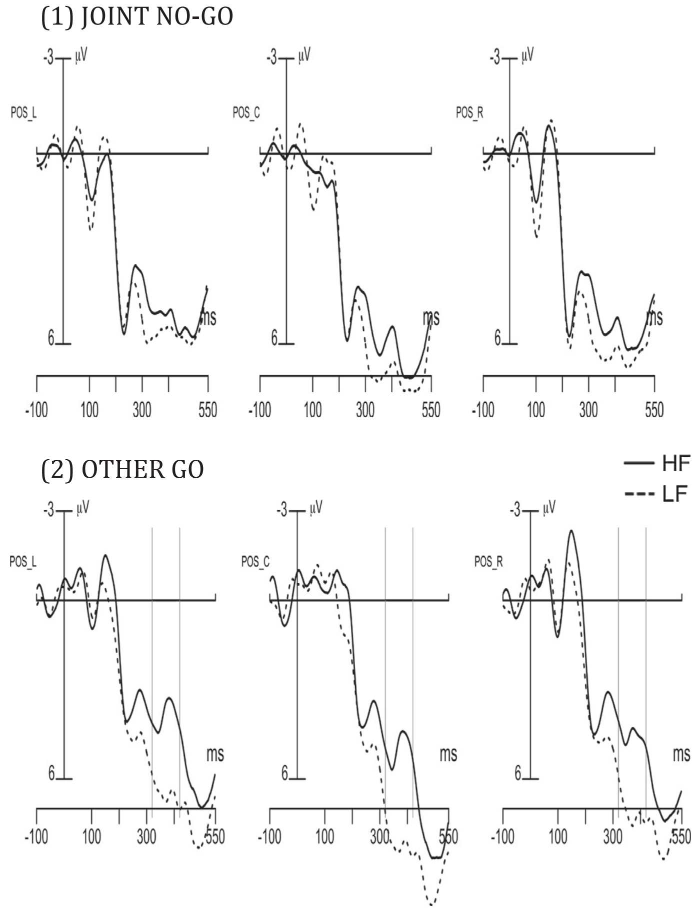

Fig. 4. EEG results for the two types of no-go trials in the JOINT condition. Grand averages and topographical analyses (P300) are presented individual no-go trials (upper panel) and other-go trials (lower panel). Solid lines represent high-frequency words (HF) and dotted lines represent low-frequency words (LF). The three figures correspond to the linear derivation of the electrodes included the three posterior regions of interest: POS_L (posterior left: T5, P3 and O1), POS_C (posterior central: Pz, PO1 and PO2) and POS_C (posterior right: T6, P4 and O2). The gray vertical lines represent the time-windows in which the frequency effect was significant.

In the time-window from 420 to 550 ms (late portion of the P300), the only significant effect was the type of trial: As predicted by earlier research, ERP amplitudes for the other-go trials were more positive than those for the joint no-go trials ($F(1,29) = 4.64$, $p < .05$, $\eta^2_p = .13$). None of the other main effects or interactions were significant (all ps > .2, except the main effect of frequency: $p = .12$).

To further confirm the frequency effects for the other-go trials, we performed an additional analysis in the P300 time-window that directly compared the individual and joint conditions. This analysis included frequency, type of no-go trials, condition (individual vs. joint), region and laterality. The results revealed more positive waveforms for the joint than for the individual condition ($F(1,29) = 11.4$, $p < .01$, $\eta^2_p = .28$). The effect of frequency was significant ($F(1,29) = 8.2$, $p < .01$, $\eta^2_p = .22$) showing overall, more positive amplitudes for those pictures with low-frequency names. Importantly, lexical frequency interacted with type of no-go trial, condition and region ($F(1.1,32.4) = 3.8$, $p = .05$, $\eta^2_p = .11$) confirming the significant effect of frequency for those no-go trials in which the confederate was asked to name the pictures (Anterior: $p = .03$, Central: $p = .002$; Posterior: $p = .001$) but not for the other no-go trials (all p's > .10).

In the late portion of the P300 (420–550 ms) only the main effects of type of trial ($F(1,29) = 4.05$, $p = .053$, $\eta^2_p = .12$) and condition ($F(1,29) = 14.4$, $p < .01$, $\eta^2_p = .33$) were significant: more positive amplitudes were observed for other-go trials than for no-go trials and more positive amplitudes in the joint condition than in the individual one. No main effect of frequency or significant interactions with other variables were observed (all p's > .12).

The presence of word frequency effects for other-go trials coupled with the absence of such effects for the joint no-go trials is perhaps the most important result of the study, since it reveals that participants engaged in lexicalization processes when the confederate was asked to name the picture. This lexicalization occurred despite the fact the participants were instructed to ignore this type of trial, and succeeded in ignoring pictures on joint no-go trials that did not require the other's response, as well as on individual no-go trials. Put differently, when participants did not have to name a picture, traces of lexical access were present as long as another person had the intention to name it.

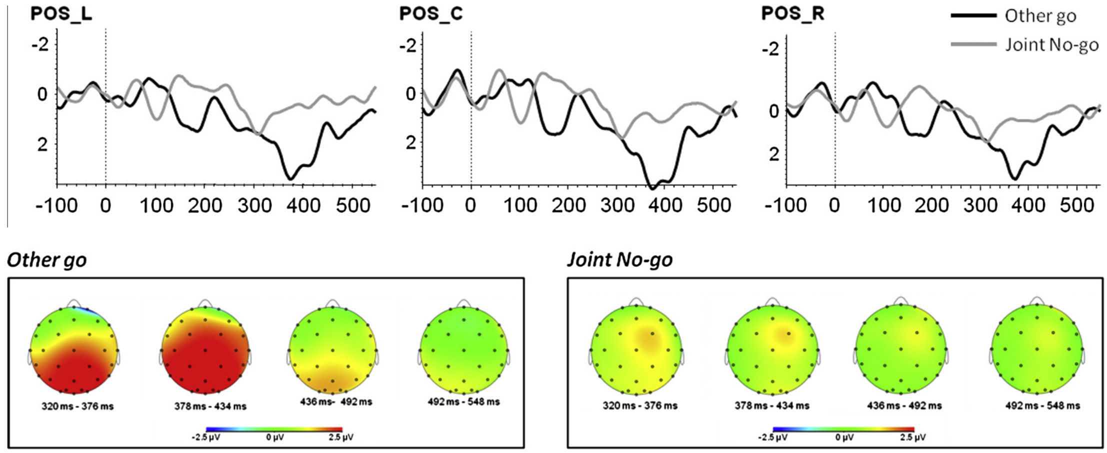

Fig. 5. Upper panel: Difference waves obtained by subtracting grand-average ERPs to high frequency words from ERPs to low-frequency words for other go (black line) and joint No-go trials (gray line) in the Joint Condition. Recording sites are posterior left POS_L (posterior left: T5, P3 and O1), Posterior central POS_C (posterior central: Pz, PO1 and PO2) and posterior right POS_R (posterior right: T6, P4 and O2). Lower panel: Topographical maps representing the frequency effect in the P300 time-window (low frequency words minus high frequency ones). Positive differences (red colors) correspond to low frequency words being more positive than high-frequency ones. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4. General discussion

In this study we explored whether people engage in similar lexicalization processes when they name pictures of objects and when a task partner is naming them. We asked participants to perform a conditional naming task, in which they sometimes had to name pictures (go trials) and sometimes had to ignore them (no-go trials). Crucially, participants did the task in an individual condition in which they were alone in the room, and in a joint condition in which a confederate participant performed a part of the task. In this joint condition, there were two types of no-go trials, trials in which neither the participants nor the confederate had to name the picture (joint no-go trials), and trials in which the participant did not have to name the picture but the confederate had to do so (other-go trials). Of particular interest for this study was participants' involvement in the different sorts of no-go trials, and specifically whether lexical processing was triggered in the other-go trials as compared to the joint no-go trials. As a proxy for lexical processing, we manipulated the word-frequency of the picture names, and we assessed whether frequency effects were present for the two types of trials. The task co-representation account predicts an effect of lexical frequency on other-go trials because it assumes that participants have formed a representation of the specific task to be performed by the other. In turn, this representation will trigger a prediction of the other's specific utterance.

The following main results were observed. First, naming latencies on go trials were faster and more accurate for high compared to low frequency words, both in the individual and joint conditions, replicating previous findings (e.g., Almeida et al., 2007; Caramazza et al., 2001; Jescheniak, Meyer, Levelt, & Specific-word, 2003; Navarrete et al., 2006; Strijkers et al., 2010). Furthermore, go trials elicited ERP word-frequency effects in the expected P200 time-window (e.g., Strijkers et al., 2011) in both conditions. In the P300 time-window, in contrast, the frequency effect remained only in the joint condition. Second, for no-go trials, word frequency effects were absent in the individual condition, as revealed by the lack of ERP frequency modulations. More interesting, however, word frequency effects emerged in the joint condition, but only for other-go trials. That is, in the joint condition word frequency effects for no-go trials were (a) absent for those no-go trials (joint no-go trials) in which none of the participants had to name the picture, and (b) present for those no-go trials in which the confederate had to name the picture. In the following, we discuss the implications of these results for go and no-go trials separately.

4.1. Naming pictures individually or in social context: Self Go trials

The time-course of the word-frequency effect for go trials is consistent with the notion that lexical processes are present around 200 ms after picture onset (Costa et al., 2009; Strijkers et al., 2010). At this early time-window, frequency effects became apparent regardless of the experimental condition in which the participant was naming the pictures, individually or jointly with the confederate participant. This observation suggests that acting together does not exert any influence on the early stages of lexical processing, at least when naming is required. These results are in line with previous observations that the onset of lexical activation during object naming is associated with a modulation of the P200 component (Aristei, Melinger, & Abdel Rahman, 2011; Costa et al., 2009; Laganaro, Valente, & Perret, 2011; Strijkers et al., 2010, 2011).

However, acting together influenced the duration of the frequency effects in the ERPs. Around 300 ms after picture onset, the frequency effects for go trials (driven by increased amplitude of the P300 for low-frequency words) were only observed in the joint condition. One possibility that could account for the observed long-lasting frequency effects is that, as suggested in the speech production literature, lexicalization processes might be prolonged in circumstances in which lexical selection processes become more cognitively demanding (Costa et al., 2009; Laganaro et al., 2011).

An indication that the joint condition might have been more difficult for the participants (or cognitively more demanding) comes from the behavioral results. Participants were slower naming pictures in the joint condition than in the individual one, while the error rate remained similar in both experimental conditions. This result is at odds with previous evidence showing no effect of the social context (Sebanz et al., 2006; Tsai et al., 2006, 2008) or a social facilitation effect, that is, faster reaction times in the joint than in the individual conditions (Aiello & Douthitt, 2001; Atmaca et al., 2011; Guerin, 1993; Sebanz et al., 2003). At present, we do not have a clear explanation as to why the influence of the social context should be different in a linguistic task. One possibility is that participants may have been more cautious in deciding whether or not to respond because speaking at the same time as another person is disruptive, and impolite in our culture (whereas in earlier studies acting at the wrong time just meant pressing a button at the same time as the partner). Alternatively (but not mutually exclusive), it is possible that the slower naming latencies in the joint condition were the result of color cues being processed differently in the individual and in the joint condition. For instance, colors in the individual condition might be represented as a dichotomous variable (my color/not my color). Conversely, the same variable could be treated as having three levels in the joint condition (my color/other's color/no-one's color), making it more cognitively demanding, and thereby delaying word retrieval. Based on the current data we cannot determine which explanation is indeed behind the observed pattern, but it opens an interesting question for future research.

4.2. Task co-representation in a picture naming task: no-go trials

While in the joint no-go trials (those in which no one named the pictures) no trace of the word frequency effect was observed, such effects were indeed present in the other-go trials, with low-frequency words eliciting a more positive waveform than high-frequency words. This result suggests that participants engaged in lexicalization processes when the confederate was asked to produce the name of a picture (other-go trials). This provides support for task co-representation accounts according to which task representations include the specific actions to be performed by a co-actor (Atmaca et al., 2011; Sebanz et al., 2005). In fact, the nature of the lexicalization processes participants engage in when naming a picture seems to be similar to that of the lexicalization when they do not name it but the confederate does, as suggested by the similarity regarding the frequency effects between the go trials and the other-go trials in the joint condition. This is consistent with the hypothesis that when performing a joint task, we predict the other's actions by constructing imitative plans at the relevant stages in the production process. In that situation, our production system would act as a forward model that runs simulations of what the other is going to say (Pickering & Garrod, 2007). However, what we cannot ensure with the present data is whether lexical representations were equally activated in the participant when naming the pictures and when the confederate named them. Predicted utterances generated by the forward model have been characterized as "impoverished" representations of the actual production representations (Pickering & Garrod, 2013; but see other authors in the same volume challenging this argument; e.g., Strijkers, Runnqvist, Costa & Holcomb, 2013).

The only observed difference between go and other-go trials was in the latency of the frequency effect. Word frequency ERP effects appeared somewhat later when the confederate named the pictures (P300) than when the participant named them (P200). As already advanced in the introduction, it is possible that this delay results from an extra involvement of cognitive processes (e.g., monitoring, inhibition) to successfully refrain the response in no-go trials (e.g., Bokura et al., 2001; Fallgatter & Strik, 1999), hence affecting the timing of the processes underlying word production. Another way of interpreting these results (that is not mutually exclusive) is that having the intention to speak proactively facilitates the engagement of lexicalization processes during object naming (Strijkers et al., 2011). That is, the speed with which lexical access starts is modulated by whether a word will be finally articulated or not. Strijkers et al. (2011) reported a delayed frequency effect (350–500 ms) when participants were not responding (no-go trials) in a go/no-go object categorization task. However, while our results nicely replicate the reported delayed frequency effect in Strijkers el al. (2011), there are important differences between this study and ours, especially regarding the nature of the no-go trials (semantic categorization vs. picture naming), that prevent us from making further interpretation of the results. Importantly, in our study joint no-go trials and other go-trials were of the exact nature from the participant's point of view (i.e., no-go trials) and therefore, the observed frequency effects can only be the result of participants selecting for further processing exclusively those objects assigned to their task-partner.

Finally, consistent with the idea that inhibitory processes might be involved at some point during action prediction, we observed that the largest difference between no-go trials that nobody named and those named by the confederate was present in the late portion of the P300 time-window (420–550 ms). The late portion of the P300 amplitude was more pronounced for those no-go trials that were named by the confederate than for those no-go trials in which nobody named the pictures. This late P300 modulation replicates and extends earlier studies that have also found an increased positivity in this time window (No-go P300) on no-go trials in joint conditions (Sebanz et al., 2006; Tsai et al., 2006, 2008) and provides clear evidence that inhibitory processes operate on trials requiring a task-partner's response. This finding is consistent with the actor and with the task co-representation account, while only the task co-representation account seems to provide a straightforward explanation of the frequency effect on other-go trials.

5. Conclusion

Taken together, our results support the claim that listeners generate predictions about speakers' utterances by relying on their own action production system (Pickering & Garrod, 2013). The frequency effect observed for those no-go trials named by the confederate suggests that participants were predicting the word their partner was intending to produce, as a consequence of co-representing the other's task (Atmaca et al., 2011; Sebanz et al., 2005; van Schie et al., 2004). This is to our knowledge the first clear empirical evidence for task co-representation, and demonstrates that individuals are able to predict what their partner aims to say. These predictive processes may facilitate all kinds of joint actions, be it walking through a narrow door or having a conversation. Future studies employing more naturalistic joint action tasks are needed to address the role of prediction in facilitating fine-grained temporal coordination, for instance, in dialogue (Stivers et al., 2009).

Acknowledgements

This research was supported by two grants from the Spanish Government, PSI2011-23033 and CONSOLIDER-INGENIO2010 CSD2007-00048, and one grant from the Catalan Government, SGR 2009-1521, and from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Cooperation grant agreement n° 613465 – AThEME). CB was supported by a post-doctoral fellowship JCI-2010-06504 from the Spanish Government and is now supported by the Intra-European Fellowship (FP7-PEOPLE-2014-IEF) of the Marie Curie Actions. FMB is supported by a pre-doctoral fellowship from the Spanish Government (FPU-2009-2013) and CM was supported by the Spanish Government (Grant Juan de la Cierva) and is now supported by the Basque Foundation for Science (IKERBASQUE) and the BCBL institution.

References

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11*(9), 1109–1116.

Aiello, J. R., & Douthitt, E. A. (2001). Social facilitation from Triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice, 5*, 163–180.

Almeida, J., Knobel, M., Finkbeiner, M., & Caramazza, A. (2007). The locus of the frequency effect in picture naming: When recognizing is not enough. *Psychonomic Bulletin & Review, 14*(6), 1177–1182.

Aristei, S., Melinger, A., & Abdel Rahman, R. (2011). Electrophysiological chronometry of semantic context effects in language production. *Journal of Cognitive Neuroscience, 23*(7), 1567–1586.

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: Sharing tasks with real and imagined co-actors. *Experimental Brain Research. Experimentelle Hirnforschung. Experimentation Cerebrale, 211*(3–4), 371–385.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin Review, 10*(2), 344–380.

Bates, D., Maechler, M., Bolker, B. (2011). lme4. R package version 0.999375-38. <http://lme4.r-forge.r-project.org/>.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, 57*, 289–300.

Bles, M., & Jansma, B. M. (2008). Phonological processing of ignored objects, an fMRI investigation. *BMC Neuroscience, 9*, 20.

Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for response inhibition in a go/NoGo task. *Clinical Neurophysiology, 112*, 2224–2232.

Brown, E. C., & Brüne, M. (2012). The role of prediction in social neuroscience. *Frontiers in Human Neuroscience, 6*, 1–27.

Caramazza, A., Bi, Y., Costa, A., & Miozzo, M. (2004). What determines the speed of lexical access: Homophone or specific-word frequency? A reply to Jescheniak et al. (2013). *Journal of Experimental Psychology: Learning, Memory and Cognition, 30*, 278–282.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1430–1450.

Clark, H. H. (Ed.). (1996). *Using language*. Cambridge: Cambridge University Press.

Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proceedings of the National Academy of Sciences, 106*(50), 21442–21446.

Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods, 37*, 665–671.

Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language and Cognitive Processes, 5*(4), 313–349.

Dittrich, K., Dolk, T., Rothe-Wulf, A., Klauer, K. C., & Prinz, W. (2013). Keys and seats: Spatial response coding underlying the joint spatial compatibility effect. *Attention, Perception & Psychophysics, 75*(8), 1725–1736.

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., & Liepelt, R. (2011). How "social" is the social Simon effect? *Frontiers in Psychology, 2*.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience, 15*(2), 399–402.

Fallgatter, A. J., & Strik, W. K. (1999). The NoGo-anteriorization as a neurophysiological standard-index for cognitive response control. *International Journal of Psychophysiology, 32*, 233–238.

Gambi, C., & Pickering, M. J. (2011). A cognitive architecture for the coordination of utterances. *Frontiers in Psychology, 2*, 275.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8*(1), 8–11.

Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*(4), 468–484.

Guerin, B. (1993). *Social facilitation*. Cambridge: Cambridge University Press.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13*, 135–145.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 824.

Jescheniak, J. D., Meyer, A. S., Levelt, W. J. M., & Specific-word, M. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 432–438.

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*(4), 463–492.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. *The Psychology of Learning and Motivation, 54*, 59–101.

Kourtis, D., Sebanz, N., & Knoblich, G. (2013). Predictive representation of other people's actions in joint action planning: An EEG study. *Social Neuroscience, 8*, 31–42.
Laganaro, M., Valente, A., & Perret, C. (2011). Time course of word production in fast and slow speakers: A high density ERP topographic study. *NeuroImage, 59*, 3881–3888.
Navarrete, E., Basagni, B., Alario, F. X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *The Quarterly Journal of Experimental Psychology, 59*(10), 1681–1690.
Obhi, S. S., & Sebanz, N. (2011). Moving together: Toward understanding the mechanisms of joint action. *Experimental Brain Research Experimentelle Hirnforschung Experimentation Cerebrale, 211*(3–4), 329–336.
Philipp, A. M., & Prinz, W. (2010). Evidence for a role of the responding agent in the joint compatibility effect. *The Quarterly Journal of Experimental Psychology, 63*(11), 2159–2171.
Pickering, M. J., & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences, 11*(3), 105–110.
Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*, 329–347.
Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2013). The R Development Core Team: nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1-104.
Ramnani, N., & Miall, R. C. (2003). A system in the human brain for predicting the actions of others. *Nature Neuroscience, 7*(1), 85–90.
Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169–192.
Sahin, N. T., Pinker, S., Cash, S. S., Schomer, D., & Halgren, E. (2009). Sequential processing of lexical, grammatical, and phonological information within broca's area. *Science, 326*(5951), 445–449.
Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition, 88*(3), B11–B21.
Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: Corepresenting stimulus-response mappings. *Journal of Experimental Psychology. Human Perception and Performance, 31*(6), 1234–1246.
Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in co-acting individuals. *Journal of Cognitive Neuroscience, 18*(5), 859–870.
Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: An integrated perspective* (pp. 31–88). Amsterdam: North-Holland.
Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174–215.
Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America, 106*(26), 10587–10592.
Strijkers, K., Baus, C., Runnqvist, E., Fitzpatrick, I., & Costa, A. (2013). The temporal dynamics of first versus second language production. *Brain and Language, 127*, 6–11.
Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex, 20*, 912–928.
Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language, 65*(4), 345–362.
Strijkers, K., Runnqvist, E., Costa, A., & Holcomb, P. (2013). The poor helping the rich: How can incomplete representations monitor complete ones? *Behavioral and Brain Sciences, 36*(04), 374–375.
Tsai, C. C., Kuo, W. J., Hung, D. L., & Tzeng, O. J. L. (2008). Action co-representation is tuned to other humans. *Journal of Cognitive Neuroscience, 20*(11), 2015–2024.
Tsai, C. C., Kuo, W. J., Jing, J. T., Hung, D. L., & Tzeng, O. J. L. (2006). A common coding framework in self-other interaction: Evidence from joint action task. *Experimental Brain Research, 175*(2), 353–362.
van Schie, H. T., Mars, R. B., Coles, M. G. H., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience, 7*(5), 549–554.
Vlainic, E., Liepelt, R., Colzato, L. S., Prinz, W., & Hommel, B. (2010). The virtual co-actor: The social Simon effect does not rely on online feedback from the other. *Frontiers in Psychology, 1*.
Welsh, T. N. (2009). When 1 + 1 = 1: The unification of independent actors revealed through joint Simon effects in crossed and uncrossed effector conditions. *Human Movement Science, 28*(6), 726–737.
Wenke, D., Atmaca, S., Holländer, A., Liepelt, R., Baess, P., & Prinz, W. (2011). What is shared in joint action? issues of co-representation, response conflict, and agent identification. *Review of Philosophy and Psychology, 2*(2), 147–172.
Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131*(3), 460.
Wingfield, A. (1968). Effects of frequency on identification and naming of objects. *The American Journal of Psychology, 81*(2), 226–234.
Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*(1431), 593–602.