The effects of utterance timing and stimulation of left prefrontal cortex on the production of referential expressions

Jennifer E. Arnold [a,*], Nazbanou Nozari [b,c]

[a] Department of Psychology and Neuroscience, University of North Carolina at Chapel Hill, United States
[b] Department of Neurology, Johns Hopkins University, United States
[c] Department of Cognitive Science, Johns Hopkins University, United States

ARTICLE INFO

ABSTRACT

We examined the relationship between the timing of utterance initiation and the choice of referring expressions, e.g., pronouns (*it*), zeros (*…and went down*), or descriptive NPs (*the pink pentagon*). We examined language production in healthy adults, and used anodal transcranial direct current stimulation (tDCS) to test the involvement of the left prefrontal cortex (PFC) in the timing of utterance production and the selection of reference forms in a discourse context. Twenty-two subjects (11 anodal, 11 sham) described fast-paced actions, e.g. *The gray oval flashes, then it moves right 2 blocks*. We only examined trials in contexts that supported pronoun/zero use. For sham participants, pronouns/zeros increased on trials with longer latencies to initiate the target utterance, and trials where the previous trial was short. We argue that both of these conditions enabled greater message pre-planning and greater discourse connect-edness: The strongest predictor of pronoun/zero usage was the presence of a connector word like *and* or *then,* which was also tended to occur on trials with longer latencies. For the anodal participants, the latency effect disappeared. PFC stimulation appeared to enable participants to produce utterances with greater discourse connectedness, even while planning incrementally.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Every time speakers refer, they must choose among various forms of reference. A shape may be referred to by a detailed description (*the pink pentagon),* a pronoun (*it),* or an elliptical (zero) construction, (*… and went down).* These choices are heavily constrained by the discourse context. Pronouns tend to refer to recently mentioned and accessible entities (Ariel, 1990, 2001; Arnold, 1998, 2008, 2010; Chafe, 1976; Givón, 1983; Gundel, Hedberg, & Zacharski, 1993), and elliptical (zero) constructions like *… and Ø moves right two blocks* are usually restricted to consecutive utterances with a repeated subject. Yet the context does not provide a categorical, inflexible constraint, and often in the same context multiple forms sound acceptable. This leads to the impression that the difference between reduced and explicit expressions is one of preference, or degree of appropriateness, leaving open many questions about the cognitive mechanisms that drive referential variation. We specifically hypothesize that variability in referential form may be related to variability in the timing of utterance initiation, which reflects the degree to which the message may be pre-planned.

The current study examines this hypothesis in two ways. Behaviorally, we investigate the relationship between the timing of utterance initiation and reference form, testing the hypothesis that reduced expressions occur more often under timing conditions that support discourse connectivity, such as greater message pre-planning. Neurally, we examine how reference form is influenced by the stimulation of the prefrontal cortex (PFC), an area shown to be involved in executive function generally, and utterance planning specifically. We tested these questions in healthy adults, and used anodal transcranial direct current stimulation (tDCS) over the left dorsolateral prefrontal cortex (L-DLPFC). Half the participants performed the task under a sham setup; results from this group established the role of planning processes in language production, in the absence of stimulation. We then examined the performance of participants under stimulation to identify ways in which stimulation changes performance.

To our knowledge, our study provides the first test of the relation between the timing of utterance initiation and reference form production. This is also the first study to examine the role of PFC on reference production with tDCS.

* Corresponding author at: UNC Chapel Hill, Dept. of Psychology, Davie Hall #337B, CB #3270, Chapel Hill, NC 27599-3270, United States.
E-mail address: jarnold@email.unc.edu (J.E. Arnold).

1.1. Why might the timing of utterance initiation affect reference form?

This study examines the time needed between observing an action and describing it verbally, which is of interest because it reflects the amount of time potentially spent pre-planning an utterance. Imagine that a speaker observes a blue square moving and says *The blue square loops around the pink triangle*. Production of this utterance requires several steps, including (a) identifying the message to be communicated, i.e. the shape that is moving, the action it is performing (looping), and the shape it is looping around, (b) selecting the words for each phrase, (c) building a syntactic structure, (d) building a phonological representation, and (e) generating a phonetic representation (e.g., Levelt, 1989). For our purposes, the critical step is message planning.

Speakers tend to plan each element of an utterance in sequence, but there is variability in how much of the message is planned before articulation begins (Konopka, 2012). Speakers may begin their utterance as soon as they identify the referent of the subject NP, for example saying *The blue square* while they figure out the action and plan the rest of the message in parallel with speaking. This approach would reflect a highly **incremental** mode of message planning and speaking. Alternatively, speakers may **preplan** a larger segment of the message, where they wait until a chunk of the message (or even the entire sentence) is planned before initiating the utterance. These two alternatives represent extremes on a continuum. Pre-planning the message does not require that linguistic formulation is also pre-planned, but message planning is at the very least a necessary condition for linguistic formulation to begin. In addition, there is known variation in the scope of verbal pre-planning (Ferreira & Swets, 2002; Griffin, 2003; Konopka, 2012; Meyer, Belke, Haecker, & Mortensen, 2007; Schriefers & Teruel, 1999; Wagner, Jescheniak, & Schriefers, 2010; Wheeldon & Lahiri, 1997).

The degree of message pre-planning is influenced by two competing pressures on speech production. On one hand, the social demands of language production induce an implicit goal of fluency (Clark & Fox Tree, 2002; Clark & Wasow, 1998). Fluency requires the speaker to plan enough of a phrase ahead of time in order to utter it without pausing. On the other hand, the social demands of language also limit the time the speaker can take to plan, in that long delays can signal that the speaker is finished, or can be perceived as nonfluent.

The intuition behind our study is that the timecourse of message planning has implications for the conceptual links between utterances in a connected discourse. When speakers pre-plan one message while articulating the previous sentence, the parallel processing may encourage conceptual ties between utterances. The reason for this conceptual overlap stems from the staged nature of language production. A message is planned conceptually, and then encoded linguistically, before it is sent to the articulation stage. While message planning and linguistic encoding may overlap, the articulatory buffer has a limited capacity of 1–2 words (Garrett, 1975; Levelt, 1989). This means each component of the message must be kept active until shortly before it is articulated. In a task like this one, where the actions determine message planning, and where the actions follow each other within seconds, the speaker may often be holding one sentence in memory while viewing the movement for the next sentence and planning it conceptually. If planning of the second sentence happens while the first sentence's message is still active, the two messages have to be active simultaneously and are more likely to be linked.

Our first hypothesis is that this conceptual integration facilitates normal processes of representing discourse relations, thus increasing the speaker's tendency to produce linguistic indicators of discourse connectedness, such as pronouns. We compared this hypothesis with two alternate possibilities. One possibility is that pre-planning the target utterance does not in fact support the use of pronouns and zeros, because it requires dual-tasking, i.e. speaking one utterance while planning the next one. This may be difficult and lead to interference between articulation and planning processes. Such interference may cause difficulty remembering the discourse context, inhibiting the production of reduced expressions, contrary to what was predicted in the first hypothesis. Another possibility is that pre-planning does indeed increase the proportion of reduced forms, but not because of conceptual integration. Rather, pre-planning may instead avoid disfluency, which tends to suppress the production of reduced expressions.

Next we describe our task and measures, before we show how our hypotheses make predictions within this task.

1.2. The moving Objects paradigm and predictor variables

Participants described visual events in the Moving Objects paradigm (Nozari, Arnold, & Thompson-Schill, 2014; Fig. 1; for a video example, see https://arnoldlab.web.unc.edu/publications/supporting-materials/arnold-nozari-2017/). Our analyses focused on the contrast between modified noun phrases like *the pink pentagon*, compared with reduced expressions (pronouns and zeros). Notably, all trials in our analysis occurred in a discourse context that supported the use of pronouns and zeros, i.e., trials on which the same shape moved on the previous trial. Thus our focus is not on the contribution of the discourse context per se, but rather on how form choice varies as a function of the timing of speech with respect to the timing of the stimulus actions.

As soon as the movement was identified, the participant could begin preparing the grammatical subject. However, the identification of the action took longer. The action durations spanned from 1200 to 4820 ms (with 200 ms between each action), leading to variation in the availability of the verb. All actions except the flash action also contained information that followed the verb (e.g. how many blocks an object moved, or which object was being jumped over), and this information also varied in how quickly it was available. Participants were encouraged to speak as normally as possible, yet the fast-paced nature of the task meant that speakers could not afford to wait, or else they would likely fall behind on subsequent trials. This typically meant that speakers began speaking before an action was finished, and were still speaking when the next action began (see Fig. 2 for an illustration).

The advantage of this task is that it creates a discourse corpus in which we control the content of each message (which corresponds to an action event), but participants are free to use wording of their choice. This provides a rich set of linguistic and timing variables, which can be analyzed in order to understand the role of planning in reference production. Our analysis examines two linguistic variables: connector use and disfluency, and four timing variables: (1) latency of target utterance; (2) action duration of target utterance; (3) latency of previous utterance; (4) action duration of previous utterance. We also examine the relation between the timing of the previous utterance and the target action onset, to test whether linguistic choices are different on trials where the action overlaps with the previous utterance. The definition and predictions of each variable are explained below.

Critical for our investigation, the onset of the action was the earliest point at which speakers could begin the process of planning any part of the message. Therefore, our first independent variable measured the latency between action onset and utterance onset, as a measure of the degree of message pre-planning on the current, target trial (IV-1 = current-latency). The latency period potentially included two component measures: (1) "overlap", i.e. overlap with the previous utterance (which occurred on 92% of the trials), and (2) "planning silence", i.e. the portion of the silence