
JOURNAL ARTICLE

Thematic role predictability and planning affect word duration

Sandra A. Zerkle, Elise C. Rosa and Jennifer E. Arnold

Department of Psychology and Neuroscience, University of North Carolina, Chapel Hill, NC, US

Corresponding author: Sandra A. Zerkle (szerkle@unc.edu)

It is known that acoustic variation is influenced by the predictability of words and the information that they represent. What is unknown is whether acoustic reduction is also influenced by the referential predictability of thematic roles. We tested this question in two production experiments, where speakers heard a sentence with goal/source arguments, e.g., “Lady Mannerly [source] gave a painting to Sir Barnes [goal],” and described a picture of a subsequent action, e.g., “Sir Barnes threw it in the closet.” We analyzed the duration of full NP descriptions used to refer to the pictured character. We found that duration was shorter for references to the goal than the source, but only in Experiment 2, where the timing of the stimuli encouraged the participant to plan their response incrementally, and not Experiment 1, where participants could pre-plan their responses. The strongest finding across both experiments was that response latency predicted duration, and latency was influenced by the predictability of thematic roles: Goal continuations had significantly shorter latencies. Together, these findings suggest that thematic role predictability does affect acoustic duration, and may be related to the time needed for utterance planning.

Keywords: predictability; prosody; acoustic reduction; duration; latency

1. Introduction

Does predictability affect acoustic prominence in spoken language? Words that are more predictable in context have been shown to be reduced in various ways, both in acoustic properties (shorter duration, less prominent) and lexical form (more pronouns). What is still unknown is whether this effect extends to all types of predictability, and what the underlying mechanism is. More generally, relatively little work has examined how semantic constraints on predictability affect language production. Predictability and expectation play a central role in many current theories of language production (Hale, 2001; Jurafsky, 1996; Kehler & Rohde, 2013; Seidenberg & MacDonald, 1999) and comprehension (Altmann & Kamide, 1999; Kooijman et al., 2005; Levy, 2008; MacDonald, 2013; Staub & Clifton, 2006; van Berkum et al., 2005), and thus it is critical to understand whether its effects are similar for multiple phenomena (e.g., the production of pronouns vs. acoustic reduction), and whether all types of predictability have the same effects. In this paper, we specifically ask whether thematic role predictability is related to both variation in production difficulty and variation in spoken word duration. Our examination of thematic roles focuses on goal/source sentences, in which a transfer-of-possession occurs between two characters. The thematic role is determined by the verb, and represents the semantic role of the participants in an event (source characters start with the object of transfer, goal characters end with the object of transfer). In example (1), Kathryn plays the role of source in both sentences, while Iris plays the role of goal.

- (1) a) *Kathryn gave the textbook to Iris, and she ... (put it on the bookshelf).*
 b) *Iris received a present from Kathryn, and she ... (opened it).*

Whenever speakers produce a referring expression, they must select an appropriate expression both lexically and prosodically. For example, speakers can use names (*Kathryn*), descriptions (*that woman*), or a reduced form, such as a pronoun (*she*). Likewise, spoken words vary in their pronunciation. Pronunciation ranges from acoustically reduced, where a speaker might say the name of a person quickly, to acoustically prominent, where the name might be drawn out in time to indicate emphasis. Such variation can be measured by characteristics such as variation in duration (short vs. long), and pitch (low vs. high, or the degree of pitch movement). Both types of reduction tend to occur when the referent has been recently mentioned, especially when it was syntactically prominent (e.g., in the grammatical position of subject in the most recent sentence; Arnold, 1998, 2010; Brennan, 1995; Chafe, 1976, 1994; Givon, 1983).

A central question about both lexical and acoustic variation is how predictability affects speakers' choices. Here we focus on this question with respect to acoustic variation. Despite solid support for the role of predictability in acoustic reduction, it is not well established what the mechanism of this effect is (Arnold & Watson, 2015). In this paper, we further explore the role of predictability on acoustic reduction in two ways. First, we seek to better understand the range of predictability manipulations that influence acoustic reduction, by testing whether thematic role predictability affects acoustic variation. Second, we seek to understand how utterance planning relates to differences in acoustic variation on spoken words, by examining the impact of utterance planning time on acoustic reduction.

1.1 Does thematic role predictability affect duration?

It is well established that the duration of spoken words is related to the predictability of the information being expressed. There are two different types of predictability that have been shown to guide reference form choice. The first is lexical predictability, which is where speakers produce certain words with shorter durations and/or greater phonological reduction when they are predictable. For example, speakers tend to use shorter pronunciations for frequent words, as well as words that are highly probable given the preceding and/or following word (Aylett & Turk, 2004; Bell et al., 2009; Frank & Jaeger, 2008; Gahl & Garnsey, 2004; Gahl et al., 2012; Jurafsky et al., 1998; Jurafsky et al., 2001). Another example comes from Lieberman (1963), where the word “nine” in the common maxim, “a stitch in time saves nine” has less emphasis (lower intelligibility) than in, “the next number you will hear is nine.” All of these studies show that lexical predictability leads to acoustic reduction, in terms of less emphasis and shorter duration for these words. Many of these findings are about lexical co-occurrence (e.g., Bell et al., 2009), but other types of contextual constraints can also influence the perceived likelihood of a word, such as knowledge of a familiar saying (Lieberman, 1963).

Another type of predictability is referential predictability, where speakers are more likely to re-mention certain entities/referents than others, making those references predictable or expected in the discourse. There's evidence that certain discourse properties confer accessibility, which guides a speaker's choice of re-mentioning an entity. Grammatical role is one example of this, where entities in the subject or first-mentioned position are more likely to be re-mentioned (Arnold, 1998). Repeated referents have reduced prominence (Lam & Watson, 2010), and given referents (i.e., those that are shared by speaker and listener, are currently foregrounded, and have contextual support) have shorter duration (Fowler & Housum, 1987). Another example from Watson et al. (2008) showed that the probability in context matters: In a tic-tac-toe game played between partners on a 3×3

grid, moves that are predictable (winning moves or blocking a winning move) are more accessible and cell number referents are spoken in a shorter duration than moves that are less predictable (“Put the blue balloon in One”). These studies show that referential predictability affects variation in acoustic prominence.

However, despite the known relation between predictability and acoustic reduction, relatively little is known about the effects of thematic roles, that is, the semantic roles of referents in the linguistic context. On one hand, we might expect the acoustic realization of names to vary, based on the semantic role that the referent played in the prior context, because thematic roles are associated with referential predictability. For example, in transfer sentences like those in (1), listeners tend to expect that the speaker will continue talking about the goal character (Iris), perhaps because it is expected that she will do something with the object she received. While it is possible to continue the story with a discussion of something that Kathryn did, statistically this pattern is less common (Arnold, 2001), and when research participants are asked to invent continuations to stories, they tend to continue with the goal (Kehler et al., 2008; Stevenson et al., 1994). That is, the goal is referentially more predictable.

On the other hand, there is reason to think that thematic role predictability may be different from other forms of predictability, because its effects have been debated. This debate concerns a different dimension of referential form: The choice between pronouns and more explicit names or descriptions. Several researchers have argued that thematic roles do not affect pronoun use (Fukumura & van Gompel, 2010; Kehler et al., 2008; Kehler & Rohde, 2013; Rohde & Kehler, 2014; Stevenson et al., 1994). These studies primarily focus on causal situations, e.g., *David scared Ana because ...*, or *Ana feared David because ...*, where one argument (here, *David*) is considered to be the more likely cause of the scaring event. When the following clause provides an explanation for the scaring event, the expected cause is more likely to be mentioned—that is, it is referentially predictable. Nevertheless, several studies have shown that speakers are no more likely to use pronouns to refer to the implicit cause than the other character, and instead speakers show a general preference to use pronouns for the subject character (Fukumura & van Gompel, 2010; Kehler et al., 2008). Given that acoustic reduction and pronoun use tend to occur in the similar discourse contexts (Ariel, 1990; Arnold, 2008; Gundel et al., 1993), these findings might predict that thematic role predictability would not affect acoustic reduction.

By contrast, mixed evidence for thematic role predictability comes from a study by Kaiser et al. (2011). They examined sentences like *Mary slapped Lisa at the zoo. As a result she...* When participants were asked to invent a continuation for the active sentences, the patient (*Lisa*) was the preferred continuation, but in passive versions (*Lisa was slapped by Mary...*), there was no clear preference. Thus, the active condition led to stronger predictability than the passive condition. They also examined the duration of names occurring in the subject position of the response, and found they were shorter in the active than the passive conditions, matching the predictability difference between conditions. But critically, this analysis collapsed across references to the subject and object. This leaves questions about whether predictability effects hold when grammatical role is controlled. While they examined agent/patient sentences, we test this question with goal/source sentences.

Yet other work suggests that speakers do indeed use pronouns more for thematic roles that are more referentially predictable. Rosa & Arnold (2017; see also Arnold, 2001) examined how speakers referred to the characters in events with transfer verbs like those in (1). Participants viewed a pair of pictures (see **Figure 1**), and heard a sentence with goal and source arguments, e.g., “Lady Mannerly [source] gave a painting to Sir Barnes [goal].” They then continued the story by describing the second picture, e.g., “He threw



Figure 1: Sample trial from Rosa and Arnold (2017). The participant heard a detective describe the first picture, and provided a description of the second picture.

it in the closet” (See **Figure 1**). Participants were found to use pronouns more often to refer to the goal character (43%) than the source character (23%). This effect occurred in addition to a general preference to use pronouns more to refer to subjects (46%) than non-subjects (22%). This demonstrates the importance of testing stimuli that present the critical thematic roles in both grammatical roles, in order to distinguish any thematic role from the general tendency to use pronouns for reference to subjects. Critically, a separate group of participants rated the goal character as more likely to be mentioned in the next sentence (71%) than the source character (29%), confirming that the goal was perceived as more referentially predictable.

In sum, findings are mixed about whether thematic roles affect the production of referential expressions. This question has received the most attention with respect to pronoun production, where the only evidence for thematic role effects comes from transfer verbs. However, there is little work on whether thematic roles affect acoustic reduction at all. Thus, our primary question is whether thematic roles affect acoustic reduction. We examine this question for transfer verbs, since evidence suggests that pronoun use is more likely for goals than sources. We do so by adopting Rosa & Arnold’s (2015) published paradigm and materials for our task.

1.2 Does planning facilitation affect duration?

A secondary goal in this study is to examine the role of message and utterance planning on duration variation, to test whether it accounts for any effect of predictability. This question is relevant in the context of a debate about whether prosodic variation stems from pragmatic rules about acoustic variation, or a tendency to use reduced expressions when production is facilitated.

1.2.1 The discourse-based selectional account

The traditional explanation for acoustic variation draws on the observation that reduced referential expressions tend to be used in discourse contexts where the referent is given and

accessible (Ariel, 1990, 1996; Arnold, 2008; Brennan, 1995; Chafe, 1976, 1994; Givon, 1983; Gundel et al., 1993). This generalization explains the preference for pronouns for accessible referents (e.g., *The diver stepped forward. She jumped into the air and dove sharply*). It also explains the preference for acoustically reduced pronunciations when the referent is recently mentioned, especially in a parallel and prominent position (Bard et al., 2004; Breen et al., 2010; Fowler & Housum, 1987; Terken & Hirschberg, 1994).

The critical idea behind this account is that prosodic forms are selected on the basis of their pragmatic appropriateness. If thematic roles affect the information status of referents, they may also result in acoustic variation. Indeed, Stevenson et al. (1994) suggest that transfer verbs elicit a natural focus on the consequences of the event, that is, the goal character, which might predict that duration of references to goals should be reduced.

1.2.2 The lexical facilitation and planning account

By contrast, recent evidence suggests that acoustic reduction may also stem from processing constraints, such that words are reduced when they are easier to plan and produce. All models of language production agree that planning to speak requires facilitating conceptual, lexical, and phonological representations prior to speaking (e.g., Dell, 1986; Garrett, 1988; Levelt, 1989). Facilitation can occur at multiple levels, including lexical retrieval, but also grammatical encoding and message planning. The speed or strength of this facilitation may affect referential form selection, providing a plausible hypothesis for how predictability may affect reference form.

When words are redundant with the context, they are easier to retrieve and are pronounced with shorter durations (e.g., Aylett & Turk, 2004; Bell et al., 2009; Gahl et al., 2012; Jurafsky et al., 2001). Bell et al. (2009) argued that the contextual probability of a word affects the speed at which it is accessed, and speakers utilize word duration as a mechanism to coordinate between planning and articulation, in order to maintain fluent delivery. Evidence for the connection between planning and word duration comes from a production study by Christodoulou (2012), who found that earlier planning led to shorter word durations. Participants described picture pairs (e.g., *toaster giraffe*), and the timing of the participant's first fixation on the second word predicted the duration of the first word. Kahn and Arnold (2012) reported evidence that word durations were shorter when a referent was predictable on the basis of visual cues, but even shorter when the word itself had been heard. This supports the idea that lexical exposure facilitates subsequent production. Lexical facilitation may also stem from referential predictability (Lam & Watson, 2010), leading to faster utterance formulation, more fluent delivery, and shorter durations.

The facilitation-based account suggests that anything that supports the facilitation of lexical representations could lead to acoustic reduction. Thus, if semantic role predictability supports lexical retrieval, it would predict greater acoustic reduction for predictable references.

1.2.3 Testing planning effects

There are two different theoretical accounts for why thematic roles might affect acoustic reduction: a) The selectional account, whereby discourse status selects for reduced or unreduced prosodic forms, and b) the facilitation/planning account. Yet testing these accounts is complicated by the fact that they are not mutually exclusive. Specifically, thematic role predictability may modulate both the ease of planning and the speaker's perception of the discourse context. If a referent is predictable, the new event may be integrated with the prior discourse context more strongly, strengthening the existing discourse representation. A more cohesive discourse representation should both impact

the speed of utterance planning, and lead to the choice of linguistic forms that mark connectivity. Consistent with this, Gillespie (2011) found that semantically integrated phrases (e.g., *The sweater with the tiny holes*) were pronounced with shorter durations than less integrated ones (*The sweater with the clean skirt*).

Further support comes from evidence people tend to produce pronouns (instead of names) more often in sentences that contain an explicit connective, like *and* or *then* (Arnold & Griffin, 2007; Arnold & Nozari, 2017). The presence of connective words reflects a stronger conceptual connection between the events being described in two sentences, and a greater ability or inclination for the speaker to treat the utterance as a part of the prior discourse context. That is, utterances without a discourse connective may be those in which participants treat the utterance as a new discourse segment (see also McCoy & Strube, 1999; Vonk et al., 1992).

In sum, the discourse-based selectional account is not inconsistent with the hypothesis that planning affects acoustic prominence. Thus, the goal of the present study is not to test between the two accounts. Rather, we more narrowly test whether utterance planning co-varies with acoustic variation, and whether both are influenced by the thematic role of the referent. We test three empirical questions: 1) Do thematic roles affect utterance planning (as measured by latency), 2) do thematic roles affect spoken word duration, and 3) do planning measures predict spoken word duration?

We test these questions in two experiments that use the picture-description task described above (Rosa & Arnold, 2015). As a measure of spoken word duration, we examine the duration of character names. As a metric of planning, we measured the latency to begin speaking, i.e., the silence prior to utterance onset. This reflects the time needed to plan the message to be communicated, and to do as much linguistic formulation as is necessary to meet the goals of fluent delivery (to the extent possible). We do not know precisely what the scope of formulation is in our task, but we assume that all message planning and at least some linguistic pre-planning occurs. We hypothesize that reference to a goal is easier to plan, based on its predictability. If so, we expect that goal references (compared to source references) will result in a shorter latency to begin speaking.

We also hypothesize that planning measures will correlate with variation in the duration of the referential expression. Prior research has suggested that in many tasks, speakers delay utterance onset in the same conditions in which they also slow down. For example, Kahn and Arnold (2015) examined the description of simple events like *The airplane rotates*, which on some conditions was preceded by a spoken prime word (“The airplane”). When the speaker heard the prime, they both initiated their description more quickly (shorter planning time) and spoke the target word more quickly. This is consistent with the hypothesis that the ease of planning the target word influences the time needed to plan the word, and also affects the speed of producing the word itself (see also Arnold & Watson, 2015; Lam & Watson, 2010; Watson et al., 2015). Thus, difficulty in retrieving a word often (but not always) results in both longer latencies and slower pronunciation time.

However, it is not always the case that an increase in planning time results in a slowing of speech articulation. In some cases, especially where the speaker has plenty of time to pre-plan an utterance, the cognitive demands associated with message and utterance planning will have been resolved prior to utterance onset, and no effect of planning difficulty will be apparent on word duration. This may explain why some studies fail to detect planning effects on word duration (Ferreira & Swets, 2002). Thus, we expect that the demands on utterance planning are most likely to have an effect when speakers are required to speak quickly, possibly while they are continuing to plan the utterance incrementally. In this situation, the speed of activating a word may have the strongest effects on the ability to articulate the word quickly, as well as the ability to plan the next word concurrently.

1.3 Experimental approach

The current experiments will examine the above questions more closely, using a naturalistic story-continuation task. In two different experiments, we measured both the duration of target names and the latency to begin speaking as a proxy for time-course of planning. Evidence suggests that speakers use pronouns more when referring to goals than sources that have been mentioned in transfer events (Rosa & Arnold, 2017). We test the hypothesis that this thematic role predictability effect will also extend to variation in acoustic reduction, such that goal continuations should have shorter durations of the target names than source continuations.

As a secondary question we also want to evaluate the role of utterance planning. This question comes in two parts: First, we predict that planning measures affect duration, because it is likely that both acoustic duration and planning time (latency to begin speaking) are tightly related. We also coded responses for the presence of a connective word (e.g., *and* or *then*), which has been known to correlate with planning and discourse cohesiveness (Arnold & Griffin, 2007; Arnold & Nozari, 2017). Second, we ask whether thematic role affects latency itself, because predictability might also be related to the time needed for a speaker to plan their utterance. Since goal continuations are more predictable, perhaps they are more accessible and therefore easier to plan, shortening the time needed before the beginning of speech. We want to assess the hypothesis that the effects of predictability might not guide duration directly, but instead might be mediated by planning time. This will give us insight into how and why predictability matters in reference form choice.

1.4 General study design

We used an experimental paradigm designed by Rosa and Arnold (2017; *jaapstimuli.unc.edu*). Participants were asked to participate in a story-telling exercise, in which they are given the role of a tabloid photographer. As background to the story, participants were told that they witnessed a murder and happened to capture pictures of the events surrounding the murder. The story has three male characters: Sir Barnes, the chauffeur, and the butler, and three female characters: Lady Mannerly, the chef, and the maid (**Figure 2**). Participants were asked to describe these pictures to a detective to help solve the crime, where the detective role is played by a researcher. This task was designed to be engaging and interesting for participants, and to encourage participants to develop a rich discourse representation of their conversation with the detective. These are properties of natural speech, which may support the effects of predictability on linguistic formulation.

The participant viewed pairs of ‘evidence photos,’ and heard the detective describe the first one, using a sentence that mentions both source and goal arguments in a transfer event (see full list of items in the Appendix). The participant’s job was to describe the second picture (see **Figure 1**), which shows a continuation that focuses on the target character, which is either the goal or the source. The detective’s sentence was manipulated (between-subjects) to control whether the target character had been mentioned as grammatical subject or non-subject in the previous sentence, e.g., “Lady Mannerly handed the picnic basket to Sir Barnes” versus “Sir Barnes took the picnic basket from Lady Mannerly.”

The storyline consisted of 53 pairs of sentences, which described actions depicted in the pairs of pictures. Critical trials (24) had two characters in the first picture, and only one in the second (indicating which character continues). This allowed us to control the content of the participant’s responses through the pictures, such that the continuation mentioned either a goal or source character. The events pictured in the target image varied, sometimes illustrating an intransitive event (“the maid laughed”),



Figure 2: Characters in the event-retelling paradigm (from left-to-right: The butler, the maid, Sir Barnes, the chef, Lady Mannerly, the chauffeur).

and sometimes illustrating a transitive event (“the chauffeur wiped off the gun”). In addition, the precise wording of the pictured event was chosen by participant (e.g., “the chef was worried” or “the chef wrung her hands”), leading to further variation in the structure of the response within and across items. The first picture in critical trials always depicted a transfer event, and this typically included a picture of the transferred object (except in three cases where the transfer was abstract, as in giving a backrub). The second picture sometimes but not always included an action with the transferred object: For goal continuations, two items had a different object, nine items had the same object, and one item had no object in the second picture; for source continuations, four items had a different object, four items had the same object, and four items had no object. The content of these events was designed to be natural and contribute to the overall murder-mystery story. Within this context, it is not surprising that the goal-continuation condition involves the transferred object more often than the source-continuation condition, because the goal character is in possession of the object at the end of the context sentence. This tendency is part and parcel of the predictability of goal characters.

Filler trials (29) had between one to three characters in both pictures. In each trial, two pictures were presented, and the detective described the first and the participant had to describe the second. All trials had to be presented in the same order for each participant in order to create a coherent story. This paradigm utilizes a typical trial-by-trial structure, while creating a naturalistic storytelling situation in which all of the utterances are related. It also allowed us to manipulate the linguistic context for the participant’s utterances.

The 24 critical trials were evenly divided between goal and source continuations (a between-items manipulation). As a control, half the trials in each condition included two characters of the same gender, and the other half included two characters of different genders. Example (2) illustrates four sample prompts and expected continuations, one in each condition.

- (2)
- 1a. **Goal/Subject:** “Sir Barnes received a painting from Lady Mannerly.”
[“Sir Barnes threw it in the closet.”]
 - 1b. **Goal/Non-Subject:** “Lady Mannerly gave a painting to Sir Barnes.”
[“Sir Barnes threw it in the closet.”]
 - 2a. **Source/Subject:** “The chauffeur handed the baskets to the chef.”
[“The chauffeur opened the door.”]
 - 2b. **Source/Non-Subject:** “The chef got the baskets from the chauffeur.”
[“The chauffeur opened the door.”]

In both experiments, participants heard the sentences and saw them depicted. In Experiment 1, both depictions were displayed on computer screens (see **Figure 1**). In Experiment 2, the first sentence was acted out on a magnet board (see **Figure 3**), and the second picture for the participant’s response was displayed on a computer screen.

The timing of the stimulus presentation differed across experiments. In Experiment 1, both the context and target pictures appeared at the start of each trial, and remained on screen while the detective described the first picture. This permitted the participant to pre-plan their response in parallel with hearing the context sentence. In Experiment 2, the target sentence did not appear until after the detective sentence, which may have led participants to begin speaking while they were still planning their response. Nevertheless, in both experiments participants previewed all the ‘evidence’ pictures before beginning the main task, which reduced the demands of interpreting the event and planning the response.

2. Experiment 1

2.1 Methods

The data analyzed for Experiment 1 are also published in Rosa & Arnold (2017), as a part of an orthogonal analysis. The current study analyzes a different subset of trials for a different purpose. As an initial examination of how thematic roles affect acoustic prominence, the current analysis only examines references to the non-subject character, since the majority of references to the subject character were produced as a pronoun, or omitted entirely.



Figure 3: Magnet board for one trial used in Experiment 2.

2.1.1 Participants

Thirty-two undergraduates completed the task for class credit in the Psychology department at the University of North Carolina, Chapel Hill. Two participants were excluded for being non-native English speakers. We excluded 9 participants who did not produce at least three names/descriptions in each of the goal/non-subject and source/non-subject conditions, leaving a total of 21 participants in this analysis.

2.1.2 Materials and Design

See General Study Design above.

2.1.3 Procedure

Participants were brought into the lab and seated at a computer, and then were shown a narrated slideshow (all materials can be found at jaapstimuli.unc.edu). The slideshow told them that they were a tabloid photographer, and described the family they had been visiting and secretly taking photographs of. It then told them that a murder occurred while they were at the house, and they were going to review the photographs they had taken to help a detective solve the crime. The participants were introduced to the characters in the pictures. Then they previewed all 53 pairs of their pictures, in order. The purpose of this preview was to familiarize the participant with the series of events. This mimics the characteristics of natural language production, in which speakers typically relate information that they already know. Then participants completed a sample trial with the experimenter. The experimenter explained that the detective, who would arrive shortly, would describe the first picture in each pair. After that, the participant should say what happened next, using the second picture as a guide.

The detective then entered the room and introduced herself. Then the audio recorder was turned on and the detective sat down at her own computer. The computers were situated back-to-back, such that the detective and photographer sat facing each other, but the computer monitors blocked their ability to see one another easily (see **Figure 4**). The detective then began the first trial, displaying the pair of pictures on both screens.

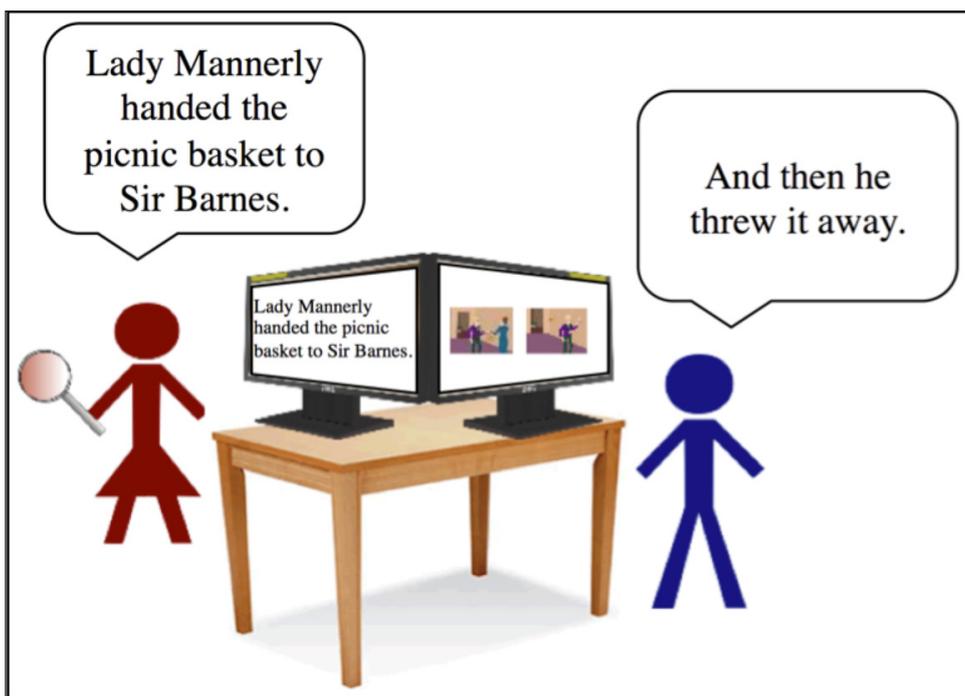


Figure 4: Experiment 1 set-up.

The detective described the first picture using a script, and then the participant said what happened next, by referring to the second picture displayed on her computer. Both pictures in the pair appeared at once on the screen for the entire duration of a trial, to encourage participants' conception of them as a coherent set. After the participant described the second picture, the detective advanced the pairs of pictures to the next trial on both screens simultaneously. A depiction of this set-up can be seen in **Figure 4** (detective is female, example participant is male). When the detective and participant had described all 53 of the events, the detective then asked the participant who had been murdered, who had committed the crime and with what weapon, and why, and other debriefing questions about the participants' familiarity with the Clue game.

2.2 Results

2.2.1 General analytical procedure

The same analytic approach was used for both experiments. Generalized linear mixed-effects models were used to account for any dependencies in the repeated measures. We used a mixed-effects linear regression (SAS 9.4 Proc Mixed) for analyses of continuous outcomes (duration and latency). All of our models include random intercepts for participants and target name. We used target name (i.e., Lady Mannerly, Sir Barnes, the butler, etc.) instead of trial as a predictor, because this allowed us to account for random effects associated with duration differences inherent to each linguistic expression. Effects coding was used for binary predictors, and is reported in each model as comparison group vs. reference group.

We used the model-building procedure outlined in Kahn & Arnold (2012), in which we first built a model with several control variables, and then retained those control variables where $|t| > 1.5$ for the final model. The purpose of this approach is to allow us to control for other relevant predictors, even if they are not of theoretical interest, while not overfitting the model. The control variables we included here were: Character gender (same vs. different), participant gender (female vs. male), referent character on right vs. left, connector use, and number of phonemes. These control models had random intercepts for participant and target name, and no random slopes to avoid overfitting/non-convergence. For each main model, the critical predictors were added and the random slopes structure was determined. Random slopes for participants and target name were included when appropriate to the design, but if any intercept or slope was estimated to be zero it wasn't included (Searle et al., 1992). The final models' fixed and random effects are reported in each model.¹

2.2.2 Response coding

Participants needed to refer to the character shown in the second picture (i.e., the target character) as the grammatical subject of their utterance for the trial to be included. The 21 participants yielded an average of 4.1 names in the goal/non-subject condition and average of 4.7 names in the source/non-subject condition. Only non-subject trials were included in the analysis, because the average number of names in the goal/subject condition was 1.8, and the average in the source/subject condition was 3.7.

Thirty trials were excluded from the final analysis, leaving 222 trials. The trials were evenly divided among goals (110 trials) and sources (112 trials). Three trials were excluded for being about non-human referents, six were excluded because the wrong character was referred to, one was excluded for using *who* as the subject, and two were excluded because

¹ We ran secondary models that included all four control predictors, and found that removing non-significant control predictors does not change the critical findings reported here.

of other mechanical issues (two trials were advanced instead of one; the picture was advanced too soon, etc.). Four trials were excluded because the participant didn't say the full name of the character (e.g., "butler" instead of "the butler," "Lany-Lady Mannerly"). An additional 14 trials were excluded due to experiment error.

Responses were also coded for use of a connective (*after*, *afterwards*, *and*, *and then*, *next*, *now*, *then*, *after that*), which was included as a control variable in the models. Use of a connective was hypothesized to indicate increased attention to the discourse context and conceptualization of the two events as a unit.

2.2.3 Audio data coding

The audio data were analyzed with Praat to measure latency to begin speaking and the duration of the target name (Boersma & Weenink, 2015). The target name was defined as the start of the first word (including "the" for characters like "the butler"), until the end of the last word. A primary undergraduate research assistant coded all the included trials, making note of four time points: The end of the beep, which signaled the presentation of both pictures for a particular trial, the end of the detective's speech (the description of the first picture), the onset of the participant's fluent speech (describing the second picture), and the length of the target character's name. A second research assistant double coded all the trials as a check of reliability. The latency measure used in this analysis is the time between the end of the detective's speech and the beginning of the participant's fluent speech. Any disfluencies said before were included in this latency measure. Trials were excluded from the latency analyses if the length was longer than 2.5 standard deviations from the mean latency; this excluded four trials. Trials were excluded from the duration analyses if the length was longer than 2.5 standard deviations from the mean duration; this excluded five trials.

2.2.4 Duration

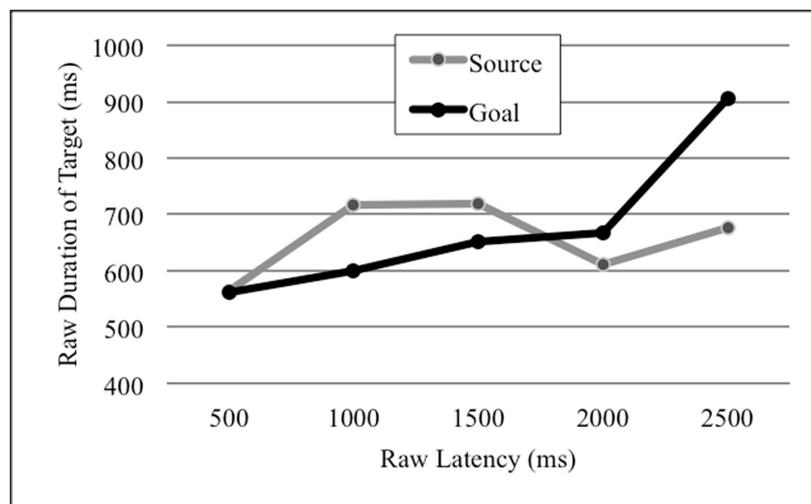
Our analyses were conducted on log-transformed measures, but here we present the raw measures for simplicity of presentation. Our first question was whether thematic role would influence the spoken duration of the character names. However, we found that it did not: The average duration for goals was 624.79 ms, and the average duration for sources was 665.09 ms. When we tested this effect in a multilevel model (with thematic role, number of phonemes, character gender, and connective use as predictors), we found no significant effect of thematic role on duration ($\beta = -0.0286$, $SE = 0.0202$, $t = -1.42$, $p = 0.158$; this model included random intercepts by participant and by expression; random slopes of thematic role were estimated to be zero by both participant and expression). In this model, number of phonemes and character gender (same vs. different) were significant predictors ($\beta = 0.04751$, $SE = 0.01055$, $t = -4.5$, $p < .0001$; $\beta = -0.0237$, $SE = 0.0103$, $t = -2.3$, $p = 0.0224$), but connective use was not ($\beta = 0.03877$, $SE = 0.03306$, $t = 1.17$, $p = 0.2422$).

Our second question was whether durations are related to measures of planning ease or difficulty. To do this, we examined whether the latency on a particular trial influenced the duration of the referring expression. As **Figure 5** shows, latency was strongly related to the duration of the target word, such that target durations were longer on trials with longer latencies. We tested the significance of this effect in a model that included thematic roles, latency, and the interaction between thematic role and latency. This model also included a measure of whether the speaker had produced a connective word (e.g., *and*, *then*), and the control variables number of phonemes and character gender. As **Table 1** shows, the only effect of interest was a significant main effect of latency. Importantly there was no significant interaction of thematic role*latency on duration, meaning the effect of latency on duration was the same for goal and source continuations (**Figure 5**).

Table 1: Experiment 1: Duration model with planning measures as predictors.

Effect	Estimate	SE	t-value	p-value
Goal vs. Source	-0.0146	0.02037	-0.72	0.4744
Log Latency	0.1973	0.05285	3.73	0.0002
Goal * Log Latency	0.04428	0.07941	0.56	0.5777
Connector Use	0.05481	0.0455	1.2	0.2297
Number of Phonemes	0.0482	0.01032	4.67	<.0001
Character Gender: Same vs. Different	-0.02607	0.009993	-2.61	0.0097

Note: This model included random intercepts for participant and expression, random slopes for connector use (by participant and by expression) and latency (by expression). Random slopes by participant for thematic role and latency were estimated to be zero, and random slope by expression for thematic role was estimated to be zero.

**Figure 5:** Experiment 1: Latency predicts duration, but no interaction with thematic role.

2.2.5 Latency

In order to understand the latency effect, we asked what contributed to latency variation. Of critical interest was whether goal continuations were initiated faster than source continuations. We found that indeed they were (**Figure 6**). The average latency for goals was 1148.12 ms, and the average latency for sources was 1415.87 ms. This effect was supported by a model of latency as the dependent measure, in which thematic role was a significant predictor (**Table 2**). This confirms our prediction that goal continuations are initiated faster than source continuations. In this model the only significant control variable was whether the target had been pictured on the left or the right in the first image. We know that participants tend to scan images left to right, so the left-hand character was likely to attract more attention than the right-hand character. Consistent with this prediction, we found that latencies were significantly shorter for left-side targets (see **Table 2**).

A potential concern (raised by the action editor for this paper) is that the difference in latency between goal and source continuations might be influenced by verb type (give-type verbs vs. receive-type verbs), and not the thematic roles themselves. Given that this analysis examines only non-subject references, the goal continuations always followed a give-type verb (e.g., *give*, *hand*), while the source continuations always followed a

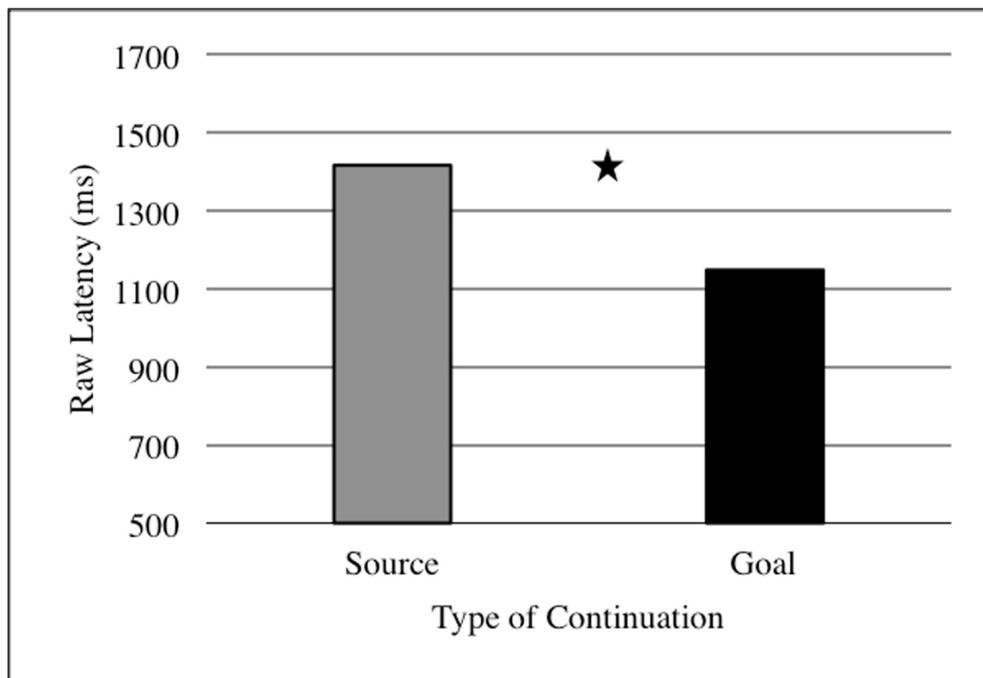


Figure 6: Experiment 1: Raw latency by type of thematic role continuation.

Table 2: Experiment 1: Latency model with goal and Referent on Right as significant predictors.

Effect	Estimate	SE	t-value	p-value
Goal vs. Source	-0.08137	0.03752	-2.17	0.0312
Referent On Right	0.06492	0.0307	2.11	0.0356

Note: This model included random intercepts for participant and expression, and random slopes for thematic role by participant and by expression. Random slope of referent on right was estimated to be zero by both participant and expression.

receive-type verb (e.g., *get*, *accept*). It is possible that the receive-type sentences are a less natural description of the picture than the corresponding give-type descriptions, and that this would increase processing difficulty. We assessed this question by asking whether the latency for the subject continuations in this task was consistent with a thematic role effect, or a verb type effect. To do so, we used the dataset for the companion study to this one (Rosa & Arnold, 2017), which contains an overlapping set of responses.² As shown in **Table 3**, goal continuations were initiated more quickly for both receive-type verbs and give-type verbs, resulting in a main effect of thematic role ($\beta = -0.12$ (.05), $t = -2.43$, $p = .0016$). There was also a non-significant effect of verb type ($\beta = -0.03$ (.03), $t = -0.97$, $p = 0.3$), and a non-significant interaction between thematic role and verb type ($\beta = 0.08$ (.05), $t = 1.52$, $p = 0.1$). This is consistent with the hypothesis that goal continuations are easier to plan, and indicates that our latency findings were not due to differences in verb type.

² As mentioned above, the analyses for these two studies sampled different subsets of the same experiment, using different criteria. The name responses from Rosa & Arnold (2017) are nearly identical to the current dataset, but they do not include data from participants who failed to produce at least 2 two pronouns, and do not exclude items on which the determiner was excluded.

Table 3: (Rosa & Arnold, 2017): Condition averages of latency to fluent speech (ms).

	Goal continuations	Source continuations
Receive-Type verbs	1201	1617
Give-Type verbs	1214	1557

2.3 Experiment 1 summary

In summary, in Experiment 1 we found that the thematic role of goal was associated with faster utterance planning, and that shorter latencies predicted shorter target name durations. However, thematic roles themselves did not have a direct effect on target word duration. Experiment 2 aimed to replicate this finding with a slightly different methodology.

3. Experiment 2

3.1 Methods

Experiment 2 used the same general experimental paradigm, with three major changes. First, the context sentence (spoken by the detective) was presented on a magnet board, where the detective acted out the scene. The motivation for this change was that it allowed us to manipulate the detective's gestures toward each character, as a way of indicating the detective's anticipation about who would be mentioned. However, this manipulation did not have any effect on our variables of interest, and thus will not be discussed in detail. Second, we expanded our analysis in this experiment to include subject references as well.

Third, the presentation of the stimulus picture (i.e., the one described by the participant) was delayed until after the detective had finished acting out the first event. This meant that participants could not begin planning their sentence until after the first sentence was finished. Even though the participant pre-viewed all the pictures (as in Experiment 1), there were too many of them to allow the participant to precisely remember the target picture on the basis of the context picture. This meant that Experiment 2 encouraged greater incremental planning than Experiment 1, which allowed us to test the impact of thematic roles and latency on word duration under conditions of incremental planning.

3.1.1 Participants

Thirty undergraduates completed the task for class credit in the Psychology department at the University of North Carolina, Chapel Hill. Six participants were excluded, leaving a total of 24 participants for whom data were analyzed. The data for this experiment were analyzed for both pronoun usage (which is reported elsewhere; Rosa & Arnold, 2017) and acoustic reduction. For this reason, we excluded any participant who did not produce at least three explicit expressions (names/descriptions) and at least three pronouns. Two participants were excluded for using fewer than three pronouns. One participant was excluded for using fewer than three names/descriptions. Two participants were excluded because the experimenter observed that they were not looking at the boards for most of the experiment (instead they looked at their computer screen). One participant was excluded because they did not view the entire background narration slideshow before meeting with the detective.

3.1.2 Materials and Design

As in Experiment 1, participants viewed pairs of pictures that were depictions of the sentence pairs described above. But in this experiment, the detective used rectangular magnet boards with magnetic pieces for each character and prop for each trial. These

boards were placed between the participant and the detective on each trial, and as the detective described the first picture, she moved the pieces to depict the action. Once the participant described the second picture, the detective moved the pieces to depict the action described. The background setting images for each trial were laminated and pasted onto 53 9" × 11" magnetic whiteboards. The characters and props were printed separately, cut out, laminated, and magnetic strips were attached to the backs of each. Small Xs were drawn onto the backgrounds to mark where the pieces should be placed at the beginning of a trial, and a second experimenter checked that all pieces were in the correct places before each participant came in. A sample board is shown in **Figure 3**. This design created some challenges, such that only a single image could be depicted for both pictures in a trial. Thus, there were slight differences on a few trials where the object of transfer changed in the pictures on the screens, but the same object was used on the magnet boards for both actions. These trials might not have been a perfect representation of the events, but the participants seemed to understand and play along accordingly.

3.1.3 Procedure

Participants were shown the same narrated slideshow as Experiment 1. The same 53 pictures were shown, and an experimenter completed a sample trial with them.

The detective then entered the room, sat down at her own computer, and the audio recorder was turned on. Like in Experiment 1, the detective's computer was angled back-to-back with the participant's computer, but in contrast with Experiment 1, the detective and participant sat so that they could still see each other and there was space on the table between them. Then the experiment would start with the second experimenter placing the first magnet board on the table, facing the participant. The detective started the slideshow on her computer and described the first picture using a script (identical to the prompt sentences given in Experiment 1), and she would slide the magnetic pieces to act out the description. The participant's screen showed a blank white screen, to encourage them to look at the detective and the board. From starting the slideshow slide of that trial, exactly three seconds pass when the detective could speak and move the pieces. After three seconds, the displays would advance automatically and the second picture appeared on the participant's screen. For the majority of the trials, we calculated latency from the onset of the stimulus picture. On six trials (five from the goal continuation condition, and one from the source continuation condition), the detective's sentence ran over the three seconds. In these trials, the latency was calculated from the end of detective speech instead of end of three seconds. When these trials were excluded, the effects in the models reported below were all the same. The participant described the second picture, and the detective would act out their description with the pieces. Then the detective advanced the pairs of pictures on both computers simultaneously, as the second experimenter removed the board and replaced it with the next. A depiction of this set-up can be seen in **Figure 7** (detective is female, participant is male, second experimenter is in a chair).

In this experiment, an additional manipulation was the detective's hand gestures while waiting for the participant to describe the second picture. Half of the participants received an anticipatory gesture always towards the subject of the previous sentence: The detective left her hand that moved the subject character on or hovering over the magnet piece until the participant finished speaking. The other half received a neutral gesture: Both of the detective's hands were taken completely off the board after acting out first sentence, waited in a neutral position while the participant described the second picture, then moved back to the board to move the pieces according to the participant's description. The anticipatory gesture was manipulated in order to test whether a speaker also takes into consideration the gestural feedback from a listener when producing an utterance, i.e.,

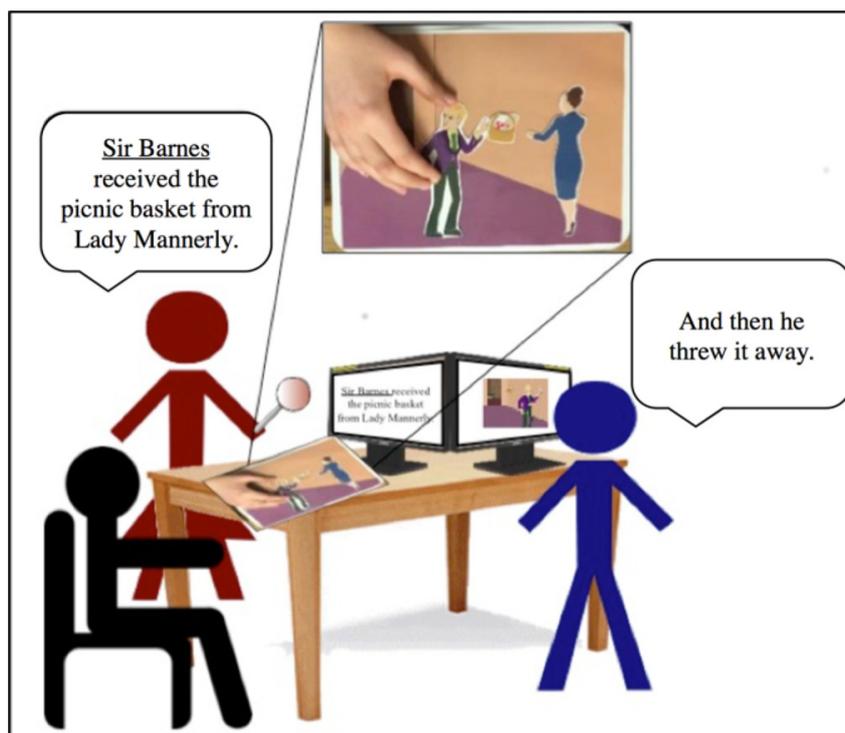


Figure 7: Experimental set-up from Experiment 2.

audience design of reference choice. In the results reported here (duration and latency), there were no effects of gesture condition.

When the detective and participant had described all the events, the detective then asked the participant who had been murdered, who had committed the crime and with what weapon, and why, and other debriefing questions about the participants' familiarity with the Clue game (similar to the questionnaire used in Experiment 1).

3.2 Results

3.2.1 General analytical procedure

We used the same analytical approach as in Experiment 1. Our dataset included all trials where the speaker used a name/description, including both subject and non-subject references.³

3.2.2 Response coding

Responses were coded in the same manner as Experiment 1. Participants needed to refer to the character pictured in the second picture of each pair as the grammatical subject of their utterance for the trial to be included.

Fifty-six trials were excluded from the final analysis, leaving 232 trials. Participants preferred to use pronouns/zeros for both subjects and goals, which meant that our final dataset was not equally distributed across conditions (subjects: $N = 71$; non-subjects: $N = 161$; goals: $N = 76$ and sources: $N = 156$). One trial was excluded for referring to the speaker ("I"), three were excluded for talking about the image ("it looks like," "is shown"), and 35 were excluded because the wrong character was referred to. Six trials were excluded because the participant didn't say the full name of the character (e.g., "chef" instead of "the chef," "Mady" instead of "Lady"), and 11 were excluded because of timing issues between trials.

³ We also analyzed the non-subject references alone, in parallel with the analysis of Experiment 1. The same effects reported here were also observed in this analysis.

As in Experiment 1, responses were also coded for use of connectives with the same coding criteria.

3.2.3 Audio data coding

The audio data were analyzed in the same manner as Experiment 1. The target name was defined as the start of the first word (including “the” for characters like “the butler”), until the end of the last word. Latency here is measured differently than in Experiment 1, due to the difference in presentation time between the two experiments. In Experiment 2, the latency measure used is from the onset of the second image on the participant’s screen to the beginning of the participant’s fluent speech (including any disfluencies, as in Experiment 1), except for the six trials on which the detective’s speech ran past the picture onset, in which case the end of the detective’s sentence was the onset of the latency measure. Durations of the target names were measured in the same way as Experiment 1. Trials were excluded from the latency analyses if the raw length was longer than three standard deviations from the mean latency; this excluded five trials. No trials were excluded from the duration analyses.

3.2.4 Duration

The critical question was whether thematic roles would influence the spoken duration of the character names. If predictability directly affects duration, then the more predictable goal continuations should have shorter duration. In this experiment, we found that it did: Target name durations in goal continuations were shorter ($M = 532.33$ ms) than in source continuations ($M = 651.53$ ms); see **Figure 8** and **Tables 4** and **5**. The control predictor number of phonemes was also significant. We also ran a model including interactions, but none were significant, so we are reporting the simpler model.

Our second question was whether planning-related measures affect duration, focusing on the effect of latency as a measure of planning time. As in Experiment 1, we found that trials with short latencies had shorter target name durations than trials with long latencies (see **Figure 9**). This resulted in a significant main effect of latency on duration (see **Table 6**). Critically, once we added latency to the duration model, the thematic role effect disappeared. We also found a significant effect of the control predictor number of phonemes. Importantly there was no significant interaction of thematic role*latency on duration, meaning the effect of latency on duration was the same for goal and source continuations (**Figure 9**).

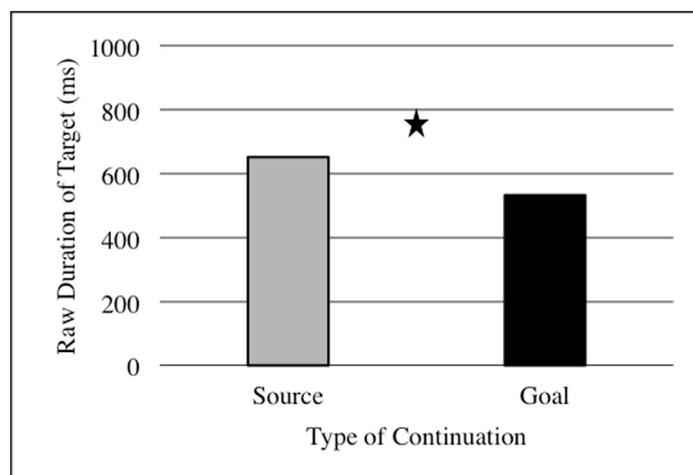


Figure 8: Experiment 2: Raw duration of target by type of thematic role continuation.

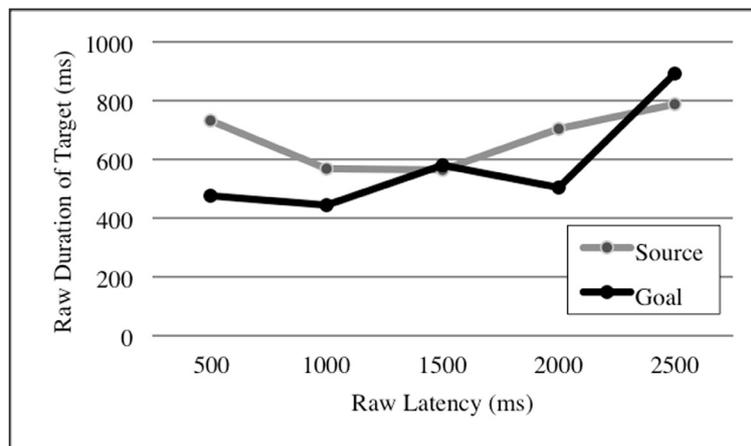
Table 4: Experiment 2: Duration of target model.

Predictor	Estimate	SE	t-value	p-value
Subject vs. Non-Subject	0.001443	0.01939	0.07	0.9408
Goal vs. Source	-0.04895	0.02232	-2.19	0.0295
Number of Phonemes	0.03706	0.01372	2.7	0.0075

Note: This model included random intercepts for participant and expression. All random slopes were estimated to be zero by both participant and by expression.

Table 5: Experiment 2: Condition averages of target duration (ms).

	Goal continuations	Source continuations
Subject continuations	520	682
Non-Subject continuations	535	635

**Figure 9:** Experiment 2: Latency predicts duration, but no interaction with thematic role.**Table 6:** Experiment 2: Duration model with planning measures as predictors.

Predictor	Estimate	SE	t-value	p-value
Subject vs. Non-Subject	-0.00000866	0.01897	0.00	1.0000
Goal vs. Source	0.261	0.2665	0.98	0.4308
Log Latency	0.1566	0.05004	3.13	0.0203
Goal * Log Latency	-0.09034	0.08411	-1.07	0.2858
Number of Phonemes	0.04599	0.01307	3.52	0.039

Note: This model included random intercepts for participant and expression, and random slopes of thematic role and latency by participant and by expression. Random slopes by subjecthood and number of phonemes estimated to be zero by both participant and by expression.

3.2.5 Latency

Our next critical question is whether thematic role affects latency itself. As in Experiment 1, in Experiment 2 there was a significant effect of thematic role on latency (**Figure 10** and **Table 7**). The average latency to begin speaking fluently was longer in trials with a source continuation (1773.60 ms) than trials with a goal continuation (1290.64 ms). This confirms our prediction that goal continuations are initiated faster than source continuations.

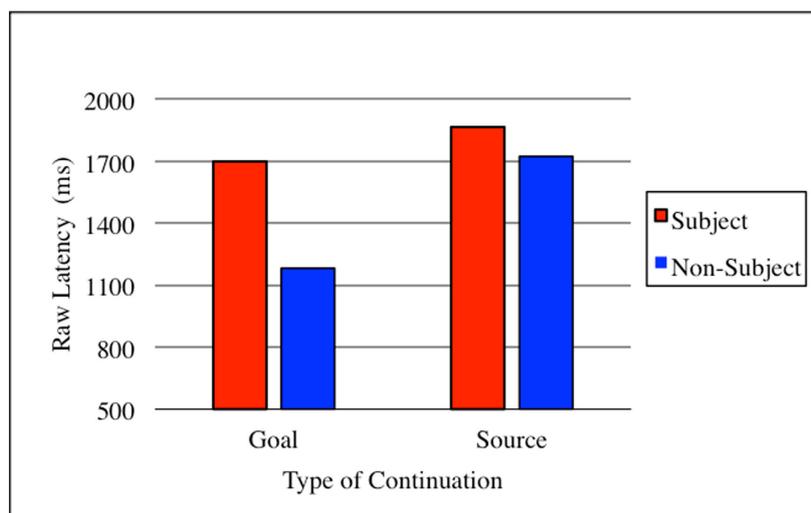


Figure 10: Experiment 2: Raw latency by type of thematic role and grammatical role continuations.

Table 7: Experiment 2: Latency model.

Effect	Estimate	SE	t-value	p-value
Subject vs. Non-Subject	0.07011	0.02592	2.7	0.0074
Goal vs. Source	-0.1664	0.02907	-5.73	<.0001
Goal * Subject	0.1634	0.05819	2.81	0.0055
Character Gender: Same vs. Different	0.03671	0.02459	1.49	0.137
Connector Use	-0.07037	0.03388	-2.08	0.0391

Note: This model included random intercepts for participant and expression. All random slopes were estimated to be zero by both participant and by expression.

We also found a main effect of subjecthood, and an interaction between thematic role and subjecthood. Surprisingly, latencies were shorter when the referent was the non-subject ($M = 1515.97$ ms) than when the referent was the subject ($M = 1827.07$ ms), despite the fact that subjects are perceived as more prominent and accessible, and latencies were especially short for non-subjects when they were the goal. This ‘reverse subjecthood’ effect may be due to the fact that this dataset includes only those trials on which the speaker used a full description. Given that pronouns are preferred for referents mentioned in subject position, the non-pronominalized trials in our dataset may have been particularly difficult. But critically, this pattern is not consistent with the alternate hypothesis that verb type drives latency, which would predict a cross-over interaction, with the longest latencies in the subject/goal and non-subject/source conditions. Condition averages are shown in **Table 8**.

3.3 Experiment 2 summary

The most notable finding from Experiment 2 was that speakers did show an effect of thematic role: References to the goal character from the stimulus sentence tended to have shorter durations than references to the source character. Thus, Experiment 2 establishes that thematic role predictability can affect spoken word duration, under at least some conditions.

Experiment 2 differed from Experiment 1 in that the stimulus picture did not appear until after the ‘detective’ had finished speaking the context sentence. This meant that the

Table 8: Experiment 2: Condition averages of latency to fluent speech (ms).

	Goal continuations	Source continuations
Subject continuations	1699	1865
Non-Subject continuations	1182	1723

participant was likely to delay utterance planning in Experiment 2. Although some pre-planning may have occurred, on the basis of the picture preview, participants likely waited for the stimulus picture to confirm the intended message. On the other hand, the social demands of conversation require a speaker to make a contribution, which puts pressure on them to begin speaking in some cases before formulation has completed (Clark & Wasow, 1998). Thus, even though our task did not explicitly give participants a deadline, there is an implicit social deadline to begin speaking before too much time had passed. This meant that participants were likely doing relatively more incremental planning than in Experiment 1. Under these circumstances, anything that facilitates message planning—such as predictability—may be critical to successful message construction and timely utterance planning.

On the other hand, the thematic role effect disappeared once we added latency to the model, and instead we observed a strong effect of latency, where short latencies led to short target durations. We also found that latency was heavily influenced by thematic role, where latencies were shorter for responses about the goal character. This finding mirrors the latency effect from Experiment 1, and supports two conclusions. First, we see strong planning effects on spoken word duration. Second, thematic role predictability affects sentence planning. We address the relationship between these two findings in the general discussion.

In our analysis of planning measures, we also examined the effect of connectives. Unlike Experiment 1, we found that connective use did predict duration, in that the presence of a connective led to longer durations. This effect is somewhat surprising, in that other studies have found that connectives co-occur with the choice of lexically reduced expressions like pronouns (Arnold & Griffin, 2007; Arnold & Nozari, 2017). In the current task, the connective may have been used as a mechanism for gaining additional time to prepare the utterance. We did not see the same connective effect on duration in Experiment 1, which is consistent with the speculation that this effect is tied to the incrementality of response planning in Experiment 2. Another possibility is that using a connective with a pronoun is more common, implying that using a connective and a name description is more rare, and thus slower.⁴

4. General discussion

The two experiments presented here provide some of the first evidence about how thematic role predictability affects acoustic reduction. Despite the fact that other measures of predictability are known to affect spoken word duration, and the fact that thematic roles are known to affect predictability, there is relatively little work on whether thematic roles affect acoustic reduction. This question is especially important given the debate about whether thematic role predictability affects pronoun production, which often occurs in similar discourse conditions to acoustic reduction (Ariel, 1990; Brennan, 1995; Gundel et al., 1993). We capitalized on recent findings that pronouns are more likely to be used to refer to goal arguments in transfer events, compared with source arguments, and

⁴ We thank an anonymous reviewer for this suggestion.

examined whether duration also varies for reference to goals and sources. Surprisingly, duration variation did not mirror pronoun variation. We did not find a consistent effect of goal/source thematic roles on duration. Instead, we found that speakers used shorter word durations for goals only in Experiment 2, when the timing of the experiment supported incremental sentence planning. Moreover, this effect went away when we added latency to the model, and instead we observed an effect of latency in both experiments: Trials with short latencies tended to have shorter target durations than trials with long latencies. Latency itself was shorter for goal continuations than source continuations. Thus, we found that thematic roles do affect acoustic reduction, but only as mediated by utterance planning. This finding makes a strong empirical contribution to the literature on acoustic reduction.

The second goal of this paper was to examine whether acoustic variation was related to utterance planning facilitation, which has been proposed as one explanation of acoustic reduction effects (Arnold & Watson, 2015). This view draws on the assumption that speech production requires planning at numerous levels. The speaker must decide on the message to be uttered, as well as formulate the linguistic elements at multiple levels. Even though some of this planning must take place before the utterance is initiated ('pre-planning'), it is well established that speakers frequently continue some of this planning during the articulation of the utterance itself ('incremental planning'; Ferreira & Swets, 2002; Levelt, 1989). Moreover, speakers vary in the proportion of the utterance that is pre-planned on any particular occasion (Ferreira & Swets, 2002). Speakers are constrained by the competing social pressures to speak reasonably fluently, but within a reasonable timeframe. In order to balance these pressures, speakers are likely to engage in some pre-planning (at least to utter the first word or words fluently), and some incremental planning. If any part of a particular response gives the speaker difficulty, it can be handled by either postponing utterance initiation (lengthening the latency), or drawing out the first words in the utterance, which are usually the target expressions in our task. Conversely, easy responses lead to speed and fluency, both in response time and speed of articulation. Such response facilitation seems especially likely for predictable information.

We tested the relationship between thematic roles and ease of planning by examining the latency to begin speaking. In both experiments, participants were faster to initiate their response when the response mentioned the goal than the source. In addition, in both experiments the latency to respond was a robust predictor of the duration of the target names. These findings together provide support for the hypothesis that word pronunciation is influenced by the ease of planning an utterance.

It seems likely that the ease of preparing sentences about goals is related to their referential predictability. Speakers are more likely to mention goals than sources in sentence-continuation tasks (Kehler et al., 2008; Stevenson et al., 1994), and corpus analysis suggests that goals are more likely to be mentioned than sources (Arnold et al., 2000). In the materials we adopted for this task, goals were also rated as more likely to be mentioned (Rosa & Arnold, 2017). We hypothesize that referential predictability may affect the ease of planning an utterance, particularly at the message level. In natural language use, predictable information is easier to retrieve from memory, enabling the speaker to plan the message more quickly. In our task, the expectation of the goal character may have facilitated the visual processing of the target slide, or it may have speeded the retrieval of the character's name. The predictability of the character in the second event may also have affected the ease of remembering the target picture from the previewed pictures. If so, in Experiment 2 this may have enabled participants to partially pre-plan the response even before the stimulus picture was available. Such an effect would be parallel to real-life situations where people relate events from memory.

The speed of message planning has direct consequences for the speed of linguistic formulation. Utterance planning is likely cascaded, such that the speaker does not need to plan the entire message before beginning linguistic formulation for some elements (Morsella & Miozzo, 2002). Thus, earlier message selection leads to faster lexical retrieval, which predicts fluent delivery and shorter word durations (e.g., Bell et al., 2009; Lam & Watson, 2010; Watson et al., 2015). In sum, our findings provide strong support for the hypothesis that acoustic reduction is influenced by message and utterance planning.

What does this say about the discourse-based selectional account of acoustic reduction? This account is based on the observation that word duration contributes to acoustic prominence or acoustic reduction, which is one linguistic cue to discourse status (Chafe, 1976; Dahan et al., 2002; Halliday, 1976). Acoustic reduction is supported by a discourse context in which the referent has a prominent status—variously described as the property of salience, accessibility, or conceptual prominence (Ariel, 1990; Arnold, 2010; Chafe, 1994; Gundel et al., 1993). In particular, previous mention in a parallel position increases the likelihood of acoustic reduction (Terken & Hirschberg, 1994). Here the question is whether referents of goal thematic roles are perceived as more prominent in the discourse context than sources. One possibility is that goal status signals discourse prominence, perhaps because of the predictability of goals. Another possibility is that referential predictability affects the strength of the speaker's memory of the discourse context, and the strength of the connections between the new event and the previous one. Predictable events will be more strongly connected with the context event in memory, and new exposures to predictable events will be more quickly integrated with the context. If so, the given status of the goal characters may be stronger than the given status of source characters, leading speakers to choose acoustically reduced expressions more often for goals. To our knowledge, no existing theories of information status propose that attention to the discourse context can vary, but this idea is consistent with other recent work from our lab (Arnold & Nozari, 2017; Zerkle & Arnold, in press).

At first glance, it may seem that our results support the planning facilitation explanation over the discourse coherence explanation. In both experiments, latency—our measure of planning time—predicted target expression duration. By contrast, the effect of thematic roles was only evident in Experiment 2, and its effect was subsumed by the latency effect. Nevertheless, we cannot rule out the discourse cohesion explanation, because the strength of the event representation may be inter-related with the time needed to engage in message planning. That is, perhaps what really matters is how strong the speaker's representation of the discourse context is, and the degree to which the relation between the events is activated, and this representation influences the speed of planning. Conversely, the strength of the discourse representation may also be influenced by the speed of utterance planning: Faster activation of a planned message may enable greater connection between that message and the previous context.

Nevertheless, our findings narrow the set of possible explanations for the relationship between thematic roles and acoustic reduction. The strong evidence for planning effects leads to three possibilities. First, it may be that only processing facilitation matters, and discourse cohesion does not. Second, it may be that discourse cohesion has two separate effects on latency and acoustic reduction. That is, if goal continuations are more cohesive than source continuations, they may be faster to plan, but for a different reason they may lead speakers to select a reduced expression as the most appropriate linguistic form for a highly accessible referent. Third, discourse cohesion may be systematically related to planning facilitation, which impacts both latency and duration.

In sum, our study provides the first evidence about whether acoustic reduction is influenced by the semantic characteristics of referents. We examined references to characters who had participated in transfer events, either as the goal or source of the transfer. We found that when speakers could not pre-plan their utterances (in Experiment 2), reference to goals were shorter than references to sources. However, this effect did not extend to a situation when speakers could pre-plan their utterances (in Experiment 1), and the effect was overshadowed by the tendency for references to be shorter when utterance initiation times were short. This provides strong support for the planning facilitation view of acoustic reduction. In addition, our findings clearly demonstrate that thematic role predictability does affect spoken word duration, but its effect is determined by its relation to predictability and utterance planning.

Additional File

The additional file for this article can be found as follows:

- **Appendix.** List of items with condition information, detective sentences, and expected continuations based on the pictures. DOI: <https://doi.org/10.5334/labphon.98.s1>

Acknowledgements

This work was funded by NSF grant 1348549 to Jennifer E. Arnold. All procedures were performed in compliance with relevant laws and institutional guidelines, and the University of North Carolina, Chapel Hill Institutional Review Board has approved them. Many thanks to Megan Fullarton, Michaela Neely, Grant Huffman, Bryan Smith, Liz Reeder, Anita Simha, Natasha Vasquez, and Taylor Beard for their help collecting and coding the data.

Competing Interests

The authors have no competing interests to declare.

References

- Altmann, G. T. and Kamide, Y. 1999. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–64. DOI: [https://doi.org/10.1016/S0010-0277\(99\)00059-1](https://doi.org/10.1016/S0010-0277(99)00059-1)
- Ariel, M. 1990. Accessing noun-phrase antecedents. London: Routledge.
- Ariel, M. 1996. Referring expressions and the +/- coreference distinction. In: Fretheim, T. and Gindell, J. K. (Eds.), *Reference and referent accessibility*, 13–35. Amsterdam: Benjamins. DOI: <https://doi.org/10.1075/pbns.38.02ari>
- Arnold, J. E. 1998. Reference Form and Discourse Patterns (Doctoral dissertation) Stanford, CA: Stanford University.
- Arnold, J. E. 2001. The effect of thematic role on pronoun use and frequency of reference continuation. *Discourse Processes*, 31, 137–162. DOI: https://doi.org/10.1207/S15326950DP3102_02
- Arnold, J. E. 2008. Reference production: Production-internal and addressee-oriented processes. *Language and Cognitive Processes*, 23(4), 495–527. DOI: <https://doi.org/10.1080/01690960801920099>
- Arnold, J. E. 2010. How speakers refer: The role of accessibility. *Language and Linguistic Compass*, 4, 187–203. DOI: <https://doi.org/10.1111/j.1749-818X.2010.00193.x>
- Arnold, J. E. and Griffin, Z. M. 2007. The effect of additional characters on choice of referring expression: Everyone counts. *Journal of memory and language*, 56(4), 521–536. DOI: <https://doi.org/10.1016/j.jml.2006.09.007>

- Arnold, J. E. and Nozari, N. 2017. The effects of utterance planning and stimulation of left prefrontal cortex on the production of referential expressions. *Cognition*, 160, 127–144. DOI: <https://doi.org/10.1016/j.cognition.2016.12.008>
- Arnold, J. E., Wasow, T., Losongco, T. and Ginstrom, R. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76(1), 28–55. DOI: <https://doi.org/10.2307/417392>
- Arnold, J. E. and Watson, D. G. 2015. Synthesizing meaning and processing approaches to prosody: Performance matters. *Language, Cognition and Neuroscience*, 30(1–2), 88–102. DOI: <https://doi.org/10.1080/01690965.2013.840733>
- Aylett, M. and Turk, A. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and speech*, 47(1), 31–56. DOI: <https://doi.org/10.1177/00238309040470010201>
- Bard, E. G., Aylett, M. P., Trueswell, J. and Tanenhaus, M. 2004. Referential form, word duration, and modeling the listener in spoken dialogue. *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions*, 173–191.
- Bell, A., Brenier, J. M., Gregory, M., Girand, C. and Jurafsky, D. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, 60(1), 92–111. DOI: <https://doi.org/10.1016/j.jml.2008.06.003>
- Boersma, P. and Weenink, D. 2015. Praat: Doing phonetics by computer [Computer program]. Version 5.4.09, retrieved March 28 2016 from: <http://www.praat.org/>.
- Breen, M., Fedorenko, E., Wagner, M. and Gibson, E. 2010. Acoustic correlates of information structure. *Language and Cognitive Processes*, 25(7–9), 1044–1098. DOI: <https://doi.org/10.1080/01690965.2010.504378>
- Brennan, S. E. 1995. Centering attention in discourse. *Language and Cognitive Processes*, 10, 137–167. DOI: <https://doi.org/10.1080/01690969508407091>
- Chafe, W. 1994. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. Chicago, IL: University of Chicago Press.
- Chafe, W. L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In: Li, C. N. (Ed.), *Subject and topic*. New York: Academic.
- Christodoulou, A. 2012. Variation in word duration and planning (Doctoral dissertation). Chapel Hill, NC: University of North Carolina.
- Clark, H. H. and Wasow, T. 1998. Repeating words in spontaneous speech. *Cognitive psychology*, 37(3), 201–242. DOI: <https://doi.org/10.1006/cogp.1998.0693>
- Dahan, D., Tanenhaus, M. K. and Chambers, C. G. 2002. Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47(2), 292–314. DOI: [https://doi.org/10.1016/S0749-596X\(02\)00001-3](https://doi.org/10.1016/S0749-596X(02)00001-3)
- Dell, G. S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological review*, 93(3), 283. DOI: <https://doi.org/10.1037/0033-295X.93.3.283>
- Ferreira, F. and Swets, B. 2002. How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46(1), 57–84. DOI: <https://doi.org/10.1006/jmla.2001.2797>
- Fowler, C. A. and Housum, J. 1987. Talkers' signalling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26, 489–504. DOI: [https://doi.org/10.1016/0749-596X\(87\)90136-7](https://doi.org/10.1016/0749-596X(87)90136-7)
- Frank, A. and Jaeger, T. F. 2008. Speaking rationally: Uniform information density as an optimal strategy for language production. In: *Proceedings of the 30th annual meeting of the cognitive science society*, 933–938. Washington, DC: Cognitive Science Society.

- Fukumura, K. and van Gompel, R. P. G. 2010. Choosing anaphoric expression: Do people take into account likelihood of reference? *Journal of Memory and Language*, 62, 52–66. DOI: <https://doi.org/10.1016/j.jml.2009.09.001>
- Gahl, S. and Garnsey, S. M. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, 80(4), 748–775. DOI: <https://doi.org/10.1353/lan.2004.0185>
- Gahl, S., Yao, Y. and Johnson, K. 2012. Why reduce? Phonological neighborhood density and phonetic reduction in spontaneous speech. *Journal of Memory and Language*, 66(4), 789–806. DOI: <https://doi.org/10.1016/j.jml.2011.11.006>
- Garrett, M. 1988. Processes in language production. In: Newmeyer, F. J. (Ed.), *Linguistics: The Cambridge survey. Vol. III. Biological and psychological aspects of language*. Cambridge, MA: Harvard University Press. DOI: <https://doi.org/10.1017/cbo9780511621062.004>
- Gillespie, M. 2011. Agreement computation in sentence production: Conceptual and temporal factors (Doctoral dissertation). Boston, MA: Northeastern University.
- Givon, T. 1983. Topic continuity in discourse: An introduction. In: Givon, T. (Ed.), *Topic continuity in discourse: A quantitative cross-language study*, 1–41. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/tsl.3>
- Gundel, J. K., Hedberg, N. and Zacharski, R. 1993. Cognitive status and the form of referring expressions in discourse. *Language*, 274–307. DOI: <https://doi.org/10.2307/416535>
- Hale, J. 2001. A probabilistic earley parser as a psycholinguistic model. *NAACL '01: Second Meeting of the North American Chapter of the Association for Computational Linguistics on Language Technologies 2001*, 1–8. DOI: <https://doi.org/10.3115/1073336.1073357>
- Halliday, M. A. 1976. System and function in language: Selected papers.
- Jurafsky, D. 1996. A probabilistic model of lexical semantic access and disambiguation. *Cognitive Science*, 20, 177–194. DOI: https://doi.org/10.1207/s15516709cog2002_1
- Jurafsky, D., Bell, A., Fosler-Lussier, E., Girand, C. and Raymond, W. D. 1998. Reduction of English function words in Switchboard. *Proceedings of ICSLP-98*, Sydney.
- Jurafsky, D., Bell, A., Gregory, M. and Raymond, W. D. 2001. Probabilistic relations between words?: Evidence from reduction in lexical production. In: Bybee, J. L. and Hopper, P. (Eds.), *Frequency and the emergence of linguistic structure*, 229–254. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/tsl.45.13jur>
- Kahn, J. M. and Arnold, J. E. 2012. A processing-centered look at the contribution of givenness to durational reduction. *Journal of Memory and Language*, 67(3), 311–325. DOI: <https://doi.org/10.1016/j.jml.2012.07.002>
- Kahn, J. M. and Arnold, J. E. 2015. Articulatory and lexical repetition effects on durational reduction: Speaker experience vs. common ground. *Language, Cognition, and Neuroscience*, 30, 103–119. DOI: <https://doi.org/10.1080/01690965.2013.848989>
- Kaiser, E., Li, D. and Holsinger, E. 2011. Exploring the lexical and acoustic consequences of referential predictability. In: Hendricks, I., Branco, A., Lalitha Devi, S. and Mitkov, R. (Eds.) *Discourse Anaphora and Anaphor Resolution Colloquium*, 171–183. Springer Berlin Heidelberg. DOI: https://doi.org/10.1007/978-3-642-25917-3_15
- Kehler, A., Kertz, L., Rohde, H. and Elman, J. L. 2008. Coherence and coreference revisited. *Journal of Semantics*, 25(1), 1–44. DOI: <https://doi.org/10.1093/jos/ffm018>
- Kehler, A. and Rohde, H. 2013. A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics*, 39(1–2), 1–37. DOI: <https://doi.org/10.1515/tl-2013-0001>

- Kooijman, V., Hagoort, P. and Cutler, A. 2005. Electrophysiological evidence for prelinguistic infants' word recognition in continuous speech. *Cognitive Brain Research*, 24(1), 109–116. DOI: <https://doi.org/10.1016/j.cogbrainres.2004.12.009>
- Lam, T. Q. and Watson, D. G. 2010. Repetition is easy: Why repeated referents have reduced prominence. *Memory & Cognition*, 38, 1137–1146. DOI: <https://doi.org/10.3758/MC.38.8.1137>
- Levelt, W. J. M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levy, R. 2008. Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. DOI: <https://doi.org/10.1016/j.cognition.2007.05.006>
- Lieberman, P. 1963. Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172–187.
- Macdonald, M. C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology*, 4(April), 226. DOI: <https://doi.org/10.3389/fpsyg.2013.00226>
- McCoy, K. and Strube, M. 1999. Generating anaphoric expressions: Pronoun or definite description? In: *Proceedings of the ACL Workshop on The Relation of Discourse/Dialogue Structure and Reference*, 63–71.
- Morsella, E. and Miozzo, M. 2002. Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 555–563. DOI: <https://doi.org/10.1037/0278-7393.28.3.555>
- Rohde, H. and Kehler, A. 2014. Grammatical and information-structural influences on pronoun production. *Language, Cognition, and Neuroscience*, 29(8), 912–927. DOI: <https://doi.org/10.1080/01690965.2013.854918>
- Rosa, E. and Arnold, J. E. 2015. Jaap Stimuli. Retrieved from: jaapstimuli.web.unc.edu.
- Rosa, E. and Arnold, J. E. 2017. Predictability affects production: Thematic roles can affect reference form selection. *Journal of Memory and Language*, 94, 43–60. DOI: <https://doi.org/10.1016/j.jml.2016.07.007>
- Searle, S., Cassella, G. and McCullough, C. 1992. *Variance Components*. New York: John Wiley-Sons. DOI: <https://doi.org/10.1002/9780470316856>
- Seidenberg, M. S. and MacDonald, M. C. 1999. A probabilistic constraints approach to language acquisition and processing. *Cognitive Science*, 23(4), 569–588. DOI: https://doi.org/10.1207/s15516709cog2304_8
- Staub, A. and Clifton, C. 2006. Syntactic prediction in language comprehension: Evidence from either ... or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 425–36. DOI: <https://doi.org/10.1037/0278-7393.32.2.425>
- Stevenson, R. J., Crawley, R. A. and Kleinman, D. 1994. Thematic roles, focus and the representation of events. *Language and Cognitive Processes*, 9(4), 519–548. DOI: <https://doi.org/10.1080/01690969408402130>
- Terken, J. and Hirschberg, J. 1994. Deaccentuation of words representing 'given' information: Effects of persistence of grammatical function and surface position. *Language and Speech*, 37(2), 125–145.
- Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V. and Hagoort, P. 2005. Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443–467. DOI: <https://doi.org/10.1037/0278-7393.31.3.443>
- Vonk, W., Hustinx, L. G. and Simons, W. H. 1992. The use of referential expressions in structuring discourse. *Language and cognitive processes*, 7(3–4), 301–333. DOI: <https://doi.org/10.1080/01690969208409389>

- Watson, D. G., Arnold, J. E. and Tanenhaus, M. K. 2008. Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 106(3), 1548–1557. DOI: <https://doi.org/10.1016/j.cognition.2007.06.009>
- Watson, D. G., Buxó-Lugo, A. and Simmons, D. C. 2015. The effect of phonological encoding on word duration: Selection takes time. In: *Explicit and Implicit Prosody in Sentence Processing*, 85–98. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-12961-7_5
- Zerkle, S. and Arnold, J. E. (in press). Discourse attention during utterance planning affects referential form choice. *Linguistics Vanguard*, 2(s1).

How to cite this article: Zerkle, S. A., Rosa, E. C. and Arnold, J. E. 2017 Thematic role predictability and planning affect word duration. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 8(1): 17, pp. 1–28, DOI: <https://doi.org/10.5334/labphon.98>

Submitted: 17 June 2017 **Accepted:** 17 June 2017 **Published:** 17 July 2017

Copyright: © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

][*Laboratory Phonology: Journal of the Association for Laboratory Phonology* is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 