First time's the charm: First-mention lengthening as an automated act

Prakaiwan Vajrabhaya and Vsevolod Kapatsinski
University of Oregon

Keywords: *Repetition, Speech production, Acoustic duration, English, Thai*

Words are longer when they are mentioned for the first time within a discourse and shorter in subsequent mentions. This Repetition Effect is usually attributed to the fact that, with repetition, the word becomes more accessible to either the speaker or the listener. In the present paper, we argue that the Repetition Effect is better seen as lengthening of words that are mentioned for the first time within a discourse (Bell et al., 2009). Furthermore, we propose that this first-mention lengthening is an automatic behavior triggered by discourse structure, rather than reflecting online changes in word accessibility for either interlocutor. In support of this proposal, we show that words are always longer when they are mentioned for the first time within a coherent stretch of discourse, even when they have previously been mentioned by the speaker to the same listener and are therefore highly accessible to both interlocutors.

Introduction

When people talk, some words are used more than once. As a word is repeated within a discourse, its duration appears to decrease, at least from the first use to the second (Anderson & Howarth, 2002; Aylett & Turk, 2004; Baker & Bradlow, 2009; Bell et al. 2009; Freeman, 2014; Fowler, 1988; Fowler et al*.,* 1997; Fowler & Housum, 1987 Galati & Brennan, 2010; Lam & Watson, 2010; Pluymaekers et al., 2005).

While the existence of this effect of repetition on duration is by now indisputable, the explanation for it remains elusive. In recognition of the theoretical disagreements, we will refer to the phenomenon as simply the *Repetition Effect.* Most popular explanations for the Repetition Effect have suggested that repetition affects the duration of a word because it affects its processing difficulty. Broadly speaking, these Processing Explanations for the Repetition Effect can be categorized into two types: Speaker-based approaches and Listener-based approaches.

According to Listener-based approaches, repetition influences duration because speakers are assumed to be sensitive to what the listener already knows or has heard before. In other words, at least on a coarse-grained level, the speaker is assumed to track the listener's state of mind and adjust word durations accordingly, making words longer when the listener might have trouble understanding them and shorter when no such difficulty is expected. (See Fowler, 1988; Frank & Jaeger, 2008; Galati & Brennan, 2010; Lindblom, 1990).

Proponents of Speaker-based approaches, on the other hand, are skeptical of the speaker's ability to model the listener, or at least of the idea that low-level dynamics of articulatory planning and execution, which manifest themselves in word duration differences, are affected by such modeling (e.g. Bard et al., 2000; Bard & Aylette, 2005; Bell et al., 2009; Bybee, 2001 and 2002). They, therefore,

attribute duration differences associated with repetition to the effect of repetition on speed of lexical access or articulatory planning within the speaker's production system. Previous mention of a word is assumed to result in a sustained increase of the word's activation level (repetition priming), making the word easier to retrieve and pronounce, which in turn results in reduced duration.

Despite their differences, processing explanations of the Repetition Effect share the assumption that duration differences among words are to be explained by how accessible those words are in the moment of production, to either the speaker or the listener. It is this assumption that we would like to question. In the study, we present data suggesting that word duration does not strictly track word accessibility and is strongly affected by discourse structure. Namely, we document that a word that is repeatedly mentioned by a speaker to the same listener can lengthen (relative to preceding productions) when the speaker starts a new story. Based on this result, we suggest that words are automatically lengthened when they are mentioned for the first time within a story.

The motivation for the automatization of first-mention lengthening is that *in the long term*, words mentioned for the first time within a coherent stretch of discourse are less accessible. As a result of automatization, words mentioned for the first time within a coherent stretch of discourse are lengthened even when *in the moment* they are highly accessible to both interlocutors. The fact that a word is mentioned for the first time within a story thus serves as a heuristic trigger for lengthening, substituting for a precise but computationally expensive real-time estimate of the word's accessibility to the current listener at the current moment (much the same way as the heuristics of Tversky & Kahneman, 1974, may substitute for computation-intensive real-time calculations of probability in decision-making).

To elaborate, Givón (1979, p.209) has claimed that grammar automates common discourse processing strategies. Whatever speakers commonly do in discourse eventually becomes conventionalized as part of the language at the community level and automatized at the level of the individual. Discourse strategies are actions that the speaker *chooses* to do in order to accomplish a conversational goal. Here, the discourse strategy of lengthening might be used to avoid misunderstanding (Lindblom, 1990). However, if the goal usually arises -- and thus a strategy is used -- in a specific linguistic environment (here, the beginning of a story), the strategy eventually becomes associated with that environment (see also Bybee, 2006). Thus, over time, the strategy becomes part of grammar: from something that the speaker *chose to do* in order to accomplish a goal, it becomes something that the speaker *cannot help but do* in a specific linguistic environment. Production of the erstwhile strategy in the context becomes automatic for the speaker, and expected for the listener. Thus, at this point lengthening is something that happens automatically at the start of a story and is conventionalized as part of how a story is told.

An effect of discourse structure on word durations has previously been reported by Fowler et al. (1997). In Experiment I, Fowler et al. had six American English speaker narrate an episode of a TV series to a listener they had just met. The narration was coded for "episode boundaries" using a variety of discourse and content cues to narrative discontinuity. Fowler et al. observed significant shortening

within the resulting episodes but not across episode boundaries, concluding that "an episode boundary blocks shortening" (p.21). In Experiment II, Fowler et al. asked 12 participants to read stories in which some paragraphs began with cues to a change of scene. Again, shortening was blocked when mentions were separated by an explicit cue to change of scene ("but in another scene") and was not blocked when the scenes formed a single narrative, the change marked by "but weeks later". Fowler et al. refer to the effect as blocking shortening rather than lengthening since they do not find significant lengthening across episode boundaries (p.21, fn.2), although they did observe numerical trends towards lengthening across episode boundaries (Figures 1-2).

Fowler et al. (1997) raised the possibility that the blocking of shortening across stories (to us, first-mention lengthening) as "the way you tell a story" in English, which is also our suggestion. However, this explanation was discarded by Fowler et al. (1997) in favor of an accessibility-based explanation. The "more fundamental" reason to reject this explanation for Fowler et al. was that "Some recent findings of our own and of other researchers lead us to doubt that durational shortening and blocking of shortening are markers provided deliberately for listeners by speakers" (Fowler et al., 1997, p.38). We share Fowler et al.'s doubts but, as noted above, we believe that much of conventionalized behavior is automatic. For example, the plural marker –s at the end of 'cats' is a conventionalized marker of plurality learned as part of acquiring English. However, the adult speaker need not deliberately choose to produce it: it is simply triggered by the semantic and morphosyntactic context. The same argument may also be made for discourse markers such as *like* in *This is, like, the most boring paper ever* (e.g. Fox Tree, 2010) that are triggered by discourse context.

The example that is, perhaps, most similar to the case under discussion is presented by definite and indefinite articles such as the English *the* and *a/an*. Like first-mention lengthening, *a/an* is used with nouns that are mentioned for the first time, while *the* is used in subsequent mentions of the same noun (e.g. Butler, 2012). The articles are also similar to first-mention lengthening in that (as we propose for first-mention lengthening) they can become associated with linguistic contexts and then used even when the following noun does not have the appropriate givenness status. For example, one almost always would use *a* in preference to *the* after *There is* in the presentational construction, even if the following noun is old information, e.g., *There is a journal that's called JML.* Here we argue that first-mention lengthening is also like the articles in being conventional and used automatically. The speaker does not intentionally and deliberately *choose* to use an article: the article is simply triggered by the appropriate discourse context. Similarly, we propose that first-mention lengthening is part of the way speakers tell stories in English and Thai without proposing that it is used strategically or deliberately. Indeed, we see the lengthening process as automatic, precisely because it is so conventional.

Fowler et al. (1997, p.38) also mention another "reason for doubting that our speakers used blocking of shortening as an optional or alternative marker of a discontinuity in a narrative is simply that it is not obvious why references to a [change in] film or scene should not have served the same purpose as the time/location markers". We would respond that a change in time/location does not

signal an end to one narrative/story and the beginning of another, while a change in the film being narrated or explicit switch of scene does. Consider the excerpts from Fowler et al.'s (1997, p.33) stimuli in (1)-(2). In (1), the narrative continuity is entirely broken by the scene-changing "episode transition": the description of the next scene has no logical connection to the preceding scene. In contrast, the narrative storyline continues fluidly through the time/location episode transition in (2). We, therefore, suggest that Fowler et al.'s data are consistent with the proposal that story/narrative boundaries trigger the speaker to lengthen words when they are mentioned for the first time within the new story.

(1) …The old woman was very happy to see the boys and asked them to stay for tea. They couldn't stay, but [Boundary][1] in another scene, you see the nanny regaling Martin with stories about Aloysius when he was a little boy at Maidenhead.

(2) "…The old woman was very happy to see the boys and asked them to stay for tea. They couldn't stay, but [Boundary] weeks later, the nanny visited them and regaled Martin with stories about Aloysius when he was a little boy at Maidenhead.

A limitation of Fowler et al.'s stimulus design is that first mention in their study is confounded with being located in the episode-initial sentence. While the first mention of a word within a story or episode does not, in general, occur in the first sentence of the story or episode, they do in Fowler et al.'s stimuli, as in (1)-(2). Episode-initial sentences have been found to elicit a higher processing load (e.g. Haberlandt et al., 1980; den Uyl & van Oostendorp, 1980). Thus, the first-mention lengthening observed by Fowler et al. (1997) can be explained by a broader version of the processing difficulty hypothesis. Instead of first-mention lengthening being driven *specifically* by lexical access difficulties (Bell et al., 2009), it could be the case that lengthening is driven by *any* kind of processing difficulty, as long as that processing difficulty competes for resources with the speaker's articulatory planning or the listener's utterance comprehension processes. Thus one aim of the present study was to see whether there is an effect of first vs. subsequent mention for words that are not part of the first sentence within a story or episode.

As previously mentioned, we view the Repetition Effect as first-mention lengthening, rather than subsequent-mention reduction. In this, we follow Bell et al. (2009). Bell et al. reasoned that the process behind the Repetition Effect should be subject to ceiling effects: a word can only be lengthened so much before becoming disfluent; it can be shortened only so much before becoming unintelligible or be understood as another word. They then noted the well-known finding that frequent words are shorter (Zipf, 1949 et seq), even when the number of segments or syllables in the word is controlled (e.g. Kapatsinski, 2010). Being shorter, frequent words have less room to reduce but more room to lengthen, when compared to infrequent words. Thus, a lengthening process should disproportionately affect frequent words, while a shortening process should be stronger with infrequent

[1] The insertion of [Boundary] is ours.

words. Bell et al. found the former: the Repetition Effect was stronger with frequent words, suggesting that it is caused by lengthening rather than shortening, which coincides with our proposal.

We should also note that reduced variants are vastly more common in discourse (e.g. Johnson, 2004; Patterson & Connine, 2001; Shuppler et al., 2011), as already implied by the fact that it is only the first mention of a word that differs in duration from the rest: all mentions of a word in discourse except the first are of similar duration (Bell et al., 2009). If the majority of mentions, the subsequent mentions, are similar in duration, and it is only the first mention that behaves differently, then it is logical to assume that it is the subsequent mentions that represent the default production target.

We do not wish to deny the influence of moment-to-moment differences in accessibility on duration. Such an in-the-moment influence could be the impetus for the automatization of story-initial lengthening, although strictly speaking it does not have to be an online influence. It could be the case that durations are adjusted post hoc based on a cue from the listener. Having produced a reduced form of a word mentioned for the first time in a story, the speaker might often encounter misunderstanding (as found by Bard et al., 2000). The speaker may then adjust future productions in the same context to be longer (cf. Norris & McQueen, 2008 for a similar argument regarding top-down influences on perception). Rather than denying a role for online processes of lengthening/shortening, we suggest that the automatized first-mention lengthening process can apply even when accessibility does not demand a deviation from the default, low-articulatory-cost reduced pronunciation.

The proposal that first mention lengthening is conventionalized implies that it is a learned behavior, rather than an inevitable outcome of production dynamics. This proposal fits well with the interaction between frequency and repetition found by Bell et al. (2009). Assuming that one has to learn first-mention lengthening, it is reasonable to assume that a speaker would be more adept at lengthening words with which s/he has had a lot of experience, including production experience (see also Erker & Guy, 2012).

In contrast, existing explanations for the Repetition Effect appear to predict the opposite interaction. Bell et al.'s (2009) own explanation of the Repetition Effect (for them, first-mention lengthening) is that the speaker lengthens words at first mention because s/he is facing lexical access difficulties. However, it is well documented that priming affects infrequent words more than it affects frequent words (the frequency attenuation effect, Forster & Davis, 1984 et seq), which overtime results in the effect of word frequency on lexical access speed being logarithmic rather than linear, e.g. Broadbent, 1967; Goldiamond & Hawkins, 1958; Howes & Solomon, 1951; Norris & McQueen, 2008; Oldfield & Wingfield, 1965).

Given the frequency attenuation effect, lexical access should be facilitated by mentioning a word to a greater extent when that word is infrequent. Since the frequency attenuation effect holds for both perception and production (frequency effects are logarithmic in both, cf. Howes & Solomon, 1951; Oldfield & Wingfield, 1965), the same problematic prediction holds whether or not the Repetition Effect is caused by accessibility of the word to the speaker or the speaker's estimate of its

accessibility to the listener. The major alternative speaker-based explanation for the Repetition Effect is that repetition causes automatization of production, thereby causing shortening (Bybee, 2001 and 2002). Again, and for the same reasons, repetition should affect low-frequency words -- whose production is not yet automatized -- more than it affects high-frequency words, whose production is highly automatic (e.g. Kapatsinski, 2010), in accordance with the Power Law of Practice (Logan, 1988; Seibel 1963). Thus, processing explanations for the Repetition Effect have trouble accounting for the results of Bell et al. (2009), leaving room for an alternative.

With the aim of evaluating the plausibility of our alternative, first mention lengthening as a conventionalized process, we created two experiments in which the same speaker produced repeated mentions of several content words to two listeners. In order to ensure generalizability, the experiments involved different languages (Thai and English)[2], different tasks (instructing the listener vs. narrating a video)[3], and different types of listener (friends of the speaker vs. confederates of the experimenter)[4]. Both experiments were designed to create opportunities for repetition within and across coherent stretches of discourse. In both experiments, we elicited crucial 'test cases' of productions in contexts where the word had previously been mentioned by the speaker and heard by the listener, yet it was being mentioned for the first time within the current stretch of discourse (see also Fowler et al., 1997). Under these circumstances, processing theories of the Repetition Effect predict shortened productions in first mention, while our automatization hypothesis suggests that productions of first mentions should be lengthened. Observing lengthened productions of words that have previously been mentioned by the current speaker to the current listener thus provides prima facie evidence for automatization of first-mention lengthening as part of discourse structure ("how you tell a story").

Experiment 1: Instruction-giving in Thai[5]

Participants

Data for this study was obtained from nine speakers of Standard/Central Thai, consisting of three female speakers and six naïve listeners. All participants were graduate students at the University of Oregon whose approximate length of stay in the United States was three years. All participants reported that they were

---

[2] Previous literature, except for Pluymaekers et al. (2005), examined the effect of repetition and reduction within a single language, English. Though Pluymaekers et al. (2005) examined Dutch, it is still a closely related language to English. Because of the limitation in languages examined, some researchers have wondered whether the reduction effect would generalize across languages (Baker & Bradlow, 2009; Ladd, 1996).

[3] cf. Bard et al. (2000) vs. Galati & Brennan (2010)

[4] Kuhlen & Brennan (2012) have cautioned that production data obtained using confederates may differ from production data obtained from using naïve listeners.

[5] The results of experiment were presented at The 17th International Congress of Phonetic Sciences, 2011, Hong Kong, China (Vajrabhaya & Kapatsinski, 2011).

active in the Thai community at the University of Oregon and were regularly exposed to Thai in both written and spoken discourse. Participants were compensated with five dollars for their participation in the study.

Procedure

Participants were asked to come to the lab in groups of three. They were friends or acquaintances who were familiar with each other. In the experiment, participants were engaged in a joint-effort spatial description task. The goal of the task was for the listener to rearrange a set of animal pictures according to the speaker's spatial description. During the task, the speaker, who was sitting across from the listener, could see the listener's board, but the listener could not see the speaker's board. All participants were explicitly informed that it was not a timed task.

There were ten pictures of common animals to be described by the speakers; all pictures had monosyllabic Thai animal names. Speakers told and retold two different "stories" involving these animal pictures to each listener. Each "story" consisted of the same set of animal picture names, but the pictures were arranged differently on the speaker's board. In other words, the *only* difference between Story 1 and Story 2 was the arrangement of the pictures. As a result, by the start of Story 2, all animal names to be mentioned had been previously mentioned during Story 1. Speakers told Story 1 (A) followed by Story 2 (B) to the first listener, and then told Story 1 (C) followed by Story 2 (D) to the second listener. There were ten-minute breaks between the tellings. The rationale for this sequence order was to avoid speakers telling the same story twice in a row to the same listener. Participants' speech was recorded using the Marantz PMD 671 digital recorder. The WAV sound files were analyzed using Praat (Boersma & Weenink, 2010).

Instances were excluded from analysis if they were adjacent to disfluencies. They were also excluded when they formed an utterance on their own because of the impossibility to calculate speech rate within the utterance independently of target word duration in such cases.

Table 1 summarizes the design of Experiment 1. In condition (A), the animal name (target word) is presented for the first time by the speaker and heard for the first time by listener 1; hence, the target word is new to both the speaker and the listener. Under any hypothesis, first mentions of words within this condition should not be reduced.

In condition (C), the word has been said many times by the speaker but is heard for the first time by listener 2; hence, the word is old to the speaker but new to listener 2. First mentions of words in this condition are expected to be reduced according to speaker-based accounts of the Repetition Effect but not according to listener-based accounts or the automatization hypothesis.

Finally, when the word is first mentioned in conditions (B) and (D), it has already been said by the speaker and heard by the listener; hence, the word is old to both interlocutors. Under processing accounts of the Repetition Effect, first mentions of words within this condition should be reduced. In contrast, the automatization hypothesis predicts that first mentions should remain unreduced even in this condition.

Table 1: The design of Experiment 1 (Spatial description task in Thai) and the Information status of the animal names to the speaker and the listener.

	To the speaker	To the listener
(A) Listener 1/Story 1	First production	First hearing
(B) Listener 1/Story 2	**Subsequent production**	**Subsequent hearing**
(C) Listener 2/Story 1	Subsequent production	First hearing
(D) Listener 2/Story 2	**Subsequent production**	**Subsequent hearing**

Coding & Analysis

The dependent variable measured was vowel duration from the release of the syllable initial-consonant until the end of vowel formants (*cf*. Abramson, 2002). The dependent variable (duration) was log transformed to achieve near-normality. The fixed effects predictors were the following:

(1) Mention within story (first mention vs. subsequent mentions); As in Bell et al. (2009), there was no significant difference between second and subsequent mentions.

(2) Story Repetition (first telling of story vs. second telling of story);

(3) Listener Repetition (first story told to listener vs. second story told to listener); and

(4) Speech rate measured as syllables per second within the utterance containing the word of interest. An utterance in this study is defined as a continuous stretch of speech delimited by pauses that are 100 ms or longer. Speech rate was measured both before and after the target word, resulting in two speech rate predictors for utterance-medial words.

Data were analyzed using linear mixed effects models in R (R Development Core Team, 2008) with the lme4 package (Bates, 2013). The model included random intercepts for speakers and items, random slopes for the fixed effects of Story Repetition, Listener Repetition, and Mention within speakers and random slopes for Story Repetition and Mention within words, resulting in near-maximal random-effects structure (Barr et al., 2013). The model did not converge if a random slope for Listener Repetition within word was included because not all words were mentioned in both stories by all speakers. Since there was no significant effect of Listener Repetition, we believe that the possible inflation of significance due to not including a random slope for the effect (Barr et al., 2013) is not an issue that would affect interpretation of the results in the present experiment. We obtained p values using the z approximation.

Results

Figure 1 shows word duration as a function of mention within story (picture arrangement) within listener. Time flows from left to right. The first four boxplots, from left to right, show duration of a word as it is repeatedly mentioned to *the first listener*. The remaining boxplots represent word duration as the words are mentioned to the second listener. Within each listener, the first two boxplots represent Story 1, while the second 2 represent Story 2. Within each story, the boxplot on the left shows first-mention durations, while the boxplot on the right shows second-mention durations.

Figure 1: Word duration boxplot. Notches show 95% confidence intervals for the median.

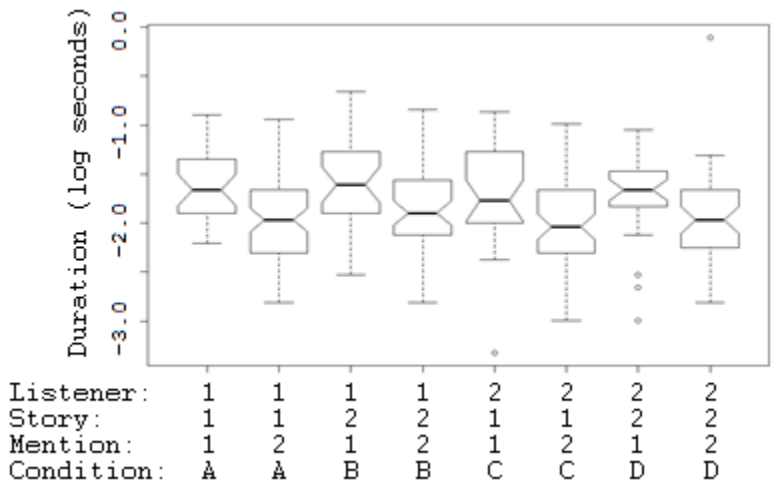

Figure 1 demonstrates that first mentions are shorter than subsequent mentions within the same story. The robustness of this effect is confirmed by mixed-effects modeling. Table 2 reports results from the mixed-effects regression model. Speech rate (syllables per second) within the same utterance before and after the word of interest was included as a covariate. Only utterance-medial occurrences were included in this analysis because there were few non-medial word occurrences and only medial occurrences allow for the measurement of speech rate both before and after the word. Table 2 shows that Repetition within Story is associated with reduced duration (second mentions within a story are shorter than first mentions): $t = -4.59$, $p < .0001$. The effects of story repetition across listeners (Listener) and listener repetition across stories (Story) did not reach significance in this analysis. There was also a significant preceding speech rate effect in the expected direction, with words being shorter when pronounced in high-speech-rate contexts: $t = -2.33$, $p = .02$.

Table 2. Fixed-effects coefficient estimates from a mixed-effects model with maximal random effects structure examining the effect of repetition on duration.

	b	*se(b)*	*t*	*P*
(Intercept)	-1.48941	0.096068	-15.504	<.0001
Repetition within story	-0.35356	0.076973	-4.593	<.0001
Repetition of story	-0.09549	0.076168	-1.254	0.21
Repetition of listener	0.07242	0.108677	0.666	0.51
Speech rate before	-0.03262	0.013997	-2.331	0.02
Speech rate after	-0.00871	0.005921	-1.471	0.14

Results remain unchanged if the analysis is re-run with only speech rate before the token or only speech rate after the token and either the 75 final or the 25 initial occurrences included. There is a significant effect of Repetition within story: b = -.32, *se(b)* = .05, t = -6.50, p < .0001 including final occurrences, b = -.36, *se(b)* = .08, t = -4.63, p < .0001 including initial ones. There is also a significant effect of the applicable Speech rate measure in both analyses: Speech rate before the word for medial and final occurrences, b = -.035, *se(b)* = .013, t = -2.76, p < .01; Speech rate after the word for medial and initial ones, b = -.012, *se(b)* = .006, t = -2.08, p = .038.

As mentioned in the introduction, Fowler et al. (1997) rejected the idea of lengthening on the basis of not observing significant lengthening across story boundaries, i.e. from subsequent mention in Story_{N} to first mention in Story_{N+1}. To determine whether there was lengthening in our data, we compared the subsequent mentions from Condition A, B and C to first mentions in Condition B, C, and D respectively. Table 3 reports the results for medial occurrences. In stark contrast to within-story repetition, repetition across a story boundary significantly *increases* word duration: the first mention of a word in a story is significantly longer than the preceding mention of that word within the preceding story. The lengthening effect remains significant if non-medial occurrences are included: b = .31, *se(b)* = .08, t = 3.85, p = .0001 with final occurrences included; b = .32, *se(b)* =.07, t = 4.35, p < .0001 with initial occurrences included. Thus, we do observe significant first-mention lengthening in our data.

Table 3. Fixed-effects coefficient estimates from a mixed-effects regression model examining the effect of story boundaries on duration.

	b	*se(b)*	*t*	*p*
(Intercept)	-1.84831	0.10722	-17.239	<.0001
Repetition across story boundary	0.31688	0.08035	3.944	<.0001
Speech rate before	-0.0341	0.01911	-1.784	0.075
Speech rate after	-0.01231	0.01144	-1.075	0.28

Discussion

In this study, duration of a word was reliably longer when it is mentioned for the first time within a story, a coherent stretch of discourse. When a speaker starts a new story, words are lengthened again for their first mention, even when the new story contains the same target words as the story that was just told to the same

listener. Thus, first-mention lengthening appears to happen even in conditions B and D when the words have been previously mentioned by the speaker to the same listener, albeit within a different story. This result is consistent with the idea that first-mention lengthening is automatized as an inalienable part of telling a story but problematic for processing-based accounts of the Repetition Effect, whether speaker-based or listener-based.

However, it is possible that the greater length at first mention and the story repetition effect could be accounted by the broader version of the processing difficulty hypothesis mentioned in the introduction. In this study, when an animal picture (target word) is first mentioned in a story, the speaker may be in the midst of planning the spatial description of the picture in relation to other pictures on the board. Even more problematically, the listener may also need time to locate the corresponding animal picture to match the speaker's spatial description. Hence, the first-mention lengthening observed in this study could have been due to a task effect. Therefore in Experiment II we decided to see if the effect would generalize to another task.

Experiment 2: Narration in American English

Participants

Data for this study was obtained from fifteen native speakers of American English, consisting of 11 female speakers and four male speakers. All participants were undergraduate students at the University of Oregon who took an introductory psychology or linguistics class that required participation in the Psychology / Linguistics Human Subjects Pool (or an alternative assignment). Participants earned research credit for their classes by taking part in the experiment.

Procedures

Participants were asked to watch a short video clip of a man making a pizza dough. They were allowed to watch the clip as many times as they wished until they were certain that they were able to tell other people how a pizza dough is made. Participants were also explicitly told that the duration of time they spent watching the clip was not being recorded, hence, allowing them to comfortably become familiar to the stimulus without time pressure. All participants watched the clip at least twice before they informed the researcher that there were ready to describe the process.

Two confederate listeners, one male and one female, were present during the experiment. The confederates were undergraduate research assistants who were in the same age range as the participants. Generally following Mol et al (2009)'s description of limited feedback, confederates' primary task was to provide non-verbal feedback (non-reactive response), which includes nodding and maintaining eye contact with the participants. Confederates limited their verbal response to occasionally saying 'uh-huh' or 'yeah'. The order in which the listeners were presented to the speakers was counterbalanced across participants. Confederates were not informed about the hypothesis.

The sequence of listeners in this study was originally used by Galati & Brennan (2010, 2014). Participants were asked to describe how a pizza dough is made to the first confederate, followed by a ten-minute break, during which

participants were asked to play a game of Tetris. When the ten-minute break was over, participants were asked to describe the same story to the second confederate, followed by another ten-minute break which participants continued the game of Tetris. Lastly, participants were asked to once again tell the same story to the first confederate. In order to avoid the implication that the first listener did not understand or follow the story on the first telling (which could result in lengthening on Listener-based accounts of the Repetition Effect), the need to retell the story was explained to the participants as resulting from a recording equipment malfunction. Specifically, the researcher told the participant that the equipment did not record properly the first time around. After the third telling, participants were debriefed and assigned research credit for their participation. In sum, participants told the same story three times, twice to the first confederate listener and once to the second confederate listener.

Participants were asked what they thought about the purpose of the experiment during debriefing. The majority reported thinking that it was a memory task. The next most popular response was that the experiment was about how we talk to men vs. women (because the confederates were of different genders). No-one mentioned catching on to the fact that the experiment was about how words are pronounced across repetitions (of the words or the story).

Participants' speech was recorded with a Marantz PMD 671 digital recorder. The WAV sound files were analyzed using Praat (Boersma & Weenink, 2010). As in Experiment 1, instances were excluded from analysis if they were adjacent to disfluencies or formed an utterance on their own. Because this study is a part of a larger study that involves gesture analysis, participants were also video-recorded while being audio-recorded.

Table 4 summarizes the experimental design. In condition (A), the information is presented and heard for the first time by listener 1; hence, the information is new to both the speaker and the listener. Thus words should be as long as they ever are when mentioned for the first time in condition (A).

In condition (B), the information is presented for the second time by the speaker but it is heard for the first time by the listener (listener 2); hence, the information is old to the speaker but new to the listener. Speaker-based processing accounts of the Repetition Effect predict that words will be reduced in this condition compared to condition (A). Listener-based accounts and the automatization hypothesis predict no reduction at first mention relative to first mention in condition (A).

In condition (C), the information is presented for the third time by the speaker and is heard for the second time by the listener; hence, the information is old to both interlocutors. Reduction is expected in this condition --relative to condition (A) – according to both speaker-based and listener-based processing explanations of the Repetition Effect. The automatization hypothesis again predicts reduction within the story but no reduction across stories.

Table 4: The design of Experiment 2 (Narrative task in American English) and the Information status of the animal names to the speaker and the listener.

	To the speaker	To the listener
(A) Listener 1	First production	First hearing
(B) Listener 2	Subsequent production	First hearing
(C) Listener 1	**Subsequent production**	**Subsequent hearing**

Bavelas & Healing (2013) and Kuhlen & Brennan (2012) raised ecological validity concerns regarding the use of confederates to elicit production data, suggesting that using data obtained from naturally occurring interaction may provide different results. We do not disagree with this statement. However, we felt that using naïve listeners (as in our Experiment 1) may be problematic for the purpose for the present study because Listener 1 would behave differently in Condition C compared to Condition A, potentially providing different feedback, which would lead the speaker to react to this feedback. The two confederates, on the other hand, would have heard the story many times prior to hearing it from the participant in the experiment, as the data from four pilot speakers (who were told that listener 1 failed to understand their instructions) are not included in the analysis. Thus, if they give any subconscious cues as to their knowledge state at all, they should give cues of having heard the story already across the three conditions, thus encouraging reduction of words when they are mentioned for the first time in all conditions. This cueing, if it occurred, would act against our automatization hypothesis, which predicts lengthening of initial mentions.

Coding & analysis

The dependent variable measured was vowel duration from the release of the syllable initial-consonant until the end of vowel formants (*cf*. Abramson, 2002). Due to the nature of the video stimulus described in the speakers' narrations, which is a man making a pizza dough, several content words, such as *pizza*, *flour*, *dough*, *knuckles*, *stretch*, and *circular*, were used repeatedly across participants. The dependent variable (duration) was square root transformed to achieve near-normality.

The fixed effects predictors were: (1) Mention within story (first mention vs. second mention vs. subsequent mentions); (2) Condition, with (A) used as the baseline level; and (3) Speech Rate measured as syllables per second within each utterance containing the word of interest (with speech rate before the word and speech rate after the word entered as separate predictors in analyzing utterance-medial durations). Data were analyzed using linear mixed effects models in R (R Development Core Team, 2008) with the lme4 package (Bates, 2013). We included

random intercepts for speakers and words and random slopes for all fixed-effects predictors of interest (excluding the speech rate covariates) within speakers and words (following Barr et al., 2013). *p* values were generated using the *z* approximation.

Results

Figure 2 shows word duration as a function of repetition within story ('Mention') across the three story repetitions ('Condition': First telling of the story to the first listener = 'A'; First telling of the story to the second listener = 'B'; Second telling of the story to the first listener = 'C') for utterance-medial word tokens. As in Experiment 1, there is a clear effect of repetition within story (confirmed to be significant using mixed-effects modeling, as shown in Table 5) with no effect of repetition across stories. In addition, again in replication of Experiment 1, there is an effect of speech rate in the expected direction.

Figure 2: Word duration boxplot. Notches show 95% confidence intervals for the median.

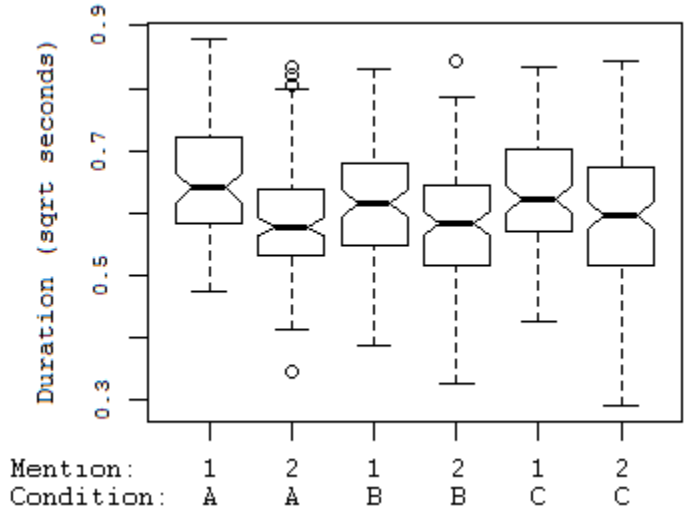

Table 5. Fixed-effects coefficient estimates from a mixed-effects model with maximal random effects structure examining the effect of repetition on duration.

	b	*se(b)*	*t*	*p*
(Intercept)	0.650348	0.023968	27.134	<.0001
Repetition within story	-0.02719	0.008065	-3.372	.0007
Condition=B	-0.01211	0.008454	-1.432	.15
Condition=C	-0.01158	0.0108	-1.072	.28
Speech rate before	-0.0051	0.001364	-3.741	.0002
Speech rate after	-0.00058	0.000584	-0.996	.32

Results again remain unchanged if boundary-adjacent instances are included. Including final occurrences, there is still a significant effect of Repetition within story (b = -.021, $se(b)$ = .007, t = -2.86, p = .004) and Speech rate (b = -.003, $se(b)$ =

.001, t = -2.68, p = .007) in addition to a significant effect of position within utterance (b = -.05, *se(b)* = .006, t = -8.10, p<.0001). This effect of position is responsible for the apparently lower median for initial mentions in Condition B in Figure 2: if position is not included in the model, words in Condition B are significantly shorter than those in Condition A. When initial instances are included, there is again a significant effect of Repetition within story (b = -.026, *se(b)* = .008, t = -3.36, p = .0008), with all other effects failing to reach significance (|t| < 1).

As in the Thai data, comparison of subsequent mentions of a word in Story_{N} to the first mentions of the same word in $\text{Story}_{\text{N+1}}$ reveals a significant lengthening effect across the story boundary, as shown in Table 6. When words that have previously mentioned are mentioned for the first time within a story, they are longer than at previous mention. The effect of story boundary remains significant if word occurrences adjacent to utterance boundaries are included: b = .022, *se(b)* = .008, t = 2.80, p = .005 including utterance-initial words, b = .021, *se(b)* = .008, t = 2.60, p = .009 including utterance-final ones).

Table 6. Fixed-effects coefficient estimates from a mixed-effects model with maximal random effects structure examining the effect of story boundaries on duration.

	b	*se(b)*	*t*	*p*
(Intercept)	0.631083	0.026738	23.602	<.0001
Repetition across story boundary	0.019365	0.008024	2.414	.016
Speech rate before	-0.00351	0.001598	-2.199	.028
Speech rate after	-0.0055	0.00239	-2.302	.021

Replicating Bell et al. (2009)'s corpus data, there was no significant duration difference between second and subsequent mentions: |t|<1 for medial words, non-final words and non-initial words. There is also no significant effect of mention if mention is entered as a continuous predictor if first mentions are excluded (|t| < 1).

Discussion

The results of the present experiment are very similar to those of Experiment 1. As in the experiment on Thai, discussed in the preceding section, first mention of a word is reliably longer than subsequent mentions within the same telling of the same story. Second mentions do not differ in duration from subsequent mentions, suggesting that it is the first mention that is the odd one out (Bell et al., 2009). When a new story starts, regardless of whether the listener has heard the story before, word duration resets when the word is first mentioned within the new story. In other words, word duration reduces within story but lengthens back to its original, first-mention duration once a new story is started. The present experiment thus confirms that the first-mention lengthening observed in Experiment 1 was not task-specific and not caused by the fact that the listener needed to locate the picture of a word's referent when it was mentioned for the first time in Experiment 1.

Bavelas & Healing (2013) and Kuhlen & Brennan (2012) raised a concern that using confederates to elicit production data and using data that is obtained from naturally occurring interaction may provide different results due to ecological

validity concerns. In this study, we find the same trend of duration resetting at story boundary across experiments even though confederates were used. Word duration resets at story boundary regardless of who is listening (even if it is somebody who has heard the story a dozen times before).

General discussion

Results from these two experiments appear to be similar despite the variability in conditions: language (Thai vs. English), sequence in which the stories were told to the two listeners and information status to the listener (1-1-2-2 vs. 1-2-1), task type (spatial description task vs. narrative task), and type of listener (naïve vs. confederate). The duration of a word becomes shorter when the word is repeated within a story: there is a significant effect of within-story repetition. However, duration of a word consistently resets (lengthens) when the speaker begins a new telling of a story and mentions the word for the first time.

These results are qualitatively consistent with prior research. Despite differing conclusions, neither Fowler et al. (1997), nor Bard et al. (2000, Experiment 1), nor Galati & Brennan (2010) found a significant decrease in word durations across stories (for words mentioned for the first time within a story). While Galati & Brennan (2010) conclude that words are reduced across stories when a story is repeated to the same listener, there was no effect of repetition across stories on duration (*minF'* = .01 for story repetition within listener in Table 2). Thus, their conclusion is based entirely on a numerical effect of story repetition within listener on clarity ratings of words mentioned for the first time within a story. The effect is not conventionally significant ($minF'(1,24) = 3.15$, $p = .08$) (Galati & Brennan, 2010, p.12). Effect of story repetition across listeners is in the unexpected direction with $p = .12$.

In contrast to the lack of significant *across*-story duration differences, Bard et al. (2000, Experiments 2-4), Bell et al. (2009), Fowler & Housum (1987), and Fowler et al. (1997) found a significant difference in duration between first and subsequent repetitions within coherent stretches of discourse.[6] In addition to replicating within-story reductions in duration, we also document that words *lengthen* relative to preceding productions by the same speaker when they are mentioned for the first time within a new story. These results hold even though words mentioned for the first time within a story generally do not occur in story-initial sentences that have been argued to be harder to process (Haberlandt et al., 1980; den Uyl & van Oostendorp, 1980): the first mention of a word occurs on average 42 seconds into the story.

We believe that the lack of reduction across stories accompanied by significant first-mention lengthening poses a problem to processing-based theories of the Repetition Effect, which propose that word duration differences correlated with repetition derive from lexical access difficulties for either the speaker (Bell et al., 2009) or the listener (Galati & Brennan, 2010; Aylett & Turk, 2004; Lindblom, 1990). Furthermore, these results also pose a problem to claims within the listener-

---

[6] Galati and Brennan (2010) did not report durations of subsequent repetitions within a story. Our results are not qualitatively changed if only first mentions within a story are analyzed.

oriented tradition that reduction/lengthening functions as a communicative signal denoting whether the word refers to given or new information (Fowler & Housum, 1987; Fowler, 1988; Chafe, 1976). If word duration reflects a word's information status, it must reflects the speaker's sensitivity to what the listener has heard vs. has not heard before (Aylett & Turk, 2004; Galati & Brennan, 2010; Lindblom, 1990). Thus, a consistent resetting pattern at story boundary, which was found in this study, should not have occurred when the word has previously been mentioned to the listener. We propose that the reset at story boundary is a result of automatic application of a conventionalized discourse strategy. Speakers reset word duration at story boundary regardless of information status and whether the listener has been exposed to the information.

To elaborate, we suggest that -- at least in adults -- first-mention lengthening is an automatized discourse-level strategy that speakers use in production. When the referent of a form has not been mentioned within the current coherent stretch of discourse, the speaker uses the full form, regardless of whether the referent is accessible to the listener. Then, within that particular stretch of coherent discourse, and only within that discourse, the full form reduces to a low-articulatory-effort default.

Since at least Clark & Marshall (1981), researchers have been concerned that modeling of the listener is a resource-demanding process that, if carried out in full detail in real time, should make conversation impossible. To alleviate this burden, several researchers in the speaker-based tradition have proposed that the speaker can default to assuming that the listener's state of mind is the same as the speaker's (e.g. Bard et al., 2000; Clark & Marshall; 1981; Horton & Keysar, 1996). In addition to this kind of egocentric processing, which generally-speaking should result in reduction of previously mentioned words, the cognitive burden of listener modeling can also be reduced by gradual incorporation of the 'generic' behavior of the listener into the structure of the discourse either over historical time at the level of the speech community or over the course of first language development within the individual.

Following Givón (1979, p.209), Grammaticalization Theory has argued that grammar results from automatization of discourse processing strategies by making them conventionalized and thus necessarily acquired to the level of unreflecting automaticity during first language acquisition. We argue that first-mention lengthening is yet another example of this automatization-through-conventionalization process. In most cases, a word mentioned for the first time within a story will be inaccessible to the listener. Thus, the use of a fuller form -- and therefore a more intelligible form, as found by Bard et al. (2000) inter alia -- is *usually* called for when a word is mentioned for the first time in a story if communication is to have a good chance of succeeding.

Once the first-mention lengthening behavior is automatized, the speaker is freed from the need to determine whether the word is in fact accessible to the listener in the moment. Production can proceed automatically and therefore effortlessly. If first mention lengthening is also conventionalized, as we propose, the listener will come to expect it even for accessible words at beginnings of stories,

thus maintaining the functional alignment between the speaker's behavior and the listener's expectations (Pickering & Garrod, 2004).

The conventionalization of goal-driven production strategies into discourse (and then grammatical) structure is quite analogous to the generalization of phonetic pressures into the conventionalized patterns of phonology. For example, there is a phonetic motivation for devoicing of stops like /b/, /d/, /g/ in utterance-final, prepausal position: vocal folds, which need to be brought together to produce voicing, are open during pauses. Hence, the speaker preparing to pause may stop voicing early in anticipation. At the same time, subglottal pressure decreases throughout the utterance, making voicing (vibration of the vocal folds due to the passage of air between them) hard to maintain at the ends of long utterances (see Myers & Padgett, in press, for a review of the literature). There is no phonetic motivation for devoicing at the ends of utterance-initial words before vowels (as in /be<u>d</u> of/). Yet, in many languages (such as German and Russian) final devoicing occurs at the ends of *all* words, even before vowels. It has been generalized beyond the context where it is motivated. Experimental work by Myers & Padgett (In press) shows that language learners are predisposed to generalize patterns from the utterance-final context to all word-final contexts. Once generalization occurs, the process is said to be phonological, and therefore conventionalized (phonology being conventionalized phonetics, e.g. Pierrehumbert, 2001). In the same way, first-mention lengthening has generalized beyond the context in which it is motivated: it occurs not only at the beginning of the first story told to the listener (onset of discourse) but at the onset of every story. This process of *domain generalization* (Myers & Padgett, in press) appears to be a sure sign of conventionalization: as a linguistic behavior becomes conventionalized, it becomes associated with smaller and smaller units (or domains).

In fact, something akin to first mention lengthening can be found as a conventionalized pattern at all levels of language, under the names of *declination* or *domain-initial strengthening*. As documented here, first mentions of words within stories are articulated more strongly than subsequent mentions. One level down, Fowler et al. (1997) document the same effect within story episodes. Further down, declination is well documented (and known to be language-specific in its magnitude) within utterances and intonational phrases (e.g. Fougeron & Keating, 1997; Pierrehumbert & Talkin, 1992). Still further down, beginnings of words are known to be articulated more strongly than ends of words (e.g. Keating et al., 2003). Finally, consonants at the beginnings of syllables are generally stronger – and longer -- than ends of syllables (e.g. Byrd, 1994; Keating et al., 1999).

These effects are often modeled as a prosodic hierarchy, with each level in the hierarchy associated with domain-initial strengthening of articulation (Fougeron & Keating, 1997; Keating et al., 2003). However, this account does not appear to be easily extended to the larger domains investigated here and in Fowler et al. (1997). Rather than positing the Prosodic Story and Prosodic Episode as levels in a prosodic hierarchy (on analogy with Prosodic Word and Phrase), we would suggest that domain-initial strengthening starts out as an online process at the discourse level and gradually undergoes domain generalization as it conventionalizes. As documented by much work within the framework of

Grammaticalization Theory, today's syntax is yesterday's discourse, today's morphology is yesterday's syntax, and even phonology is sometimes yesterday's morphology (Givón, 1979 et seq; see Bybee, 2006; and DeLancey 2011 for recent reviews). Thus, it seems profitable to consider the possibility that the domain-initial strengthening effects found at utterance, phrase, word, and maybe even syllable levels are the grammaticalized outcomes of domain generalization accompanying the conventionalization of the discourse pressure towards story-initial lengthening and, more generally, clear production.

We believe that viewing the Repetition Effect as conventionalized first-mention lengthening as opposed to online reduction caused by repetition provides a superior account of findings from prior research. Given the fact that first mention is significantly longer than all subsequent mentions, which do not differ significantly from each other (Bell et al., 2009), the 'reduced' form is much more common in production than the 'full' or lengthened form. Thus, it appears reasonable to assume that the reduced form constitutes the default production plan that is departed from only when such departures are 'forced' by the threat of listener misunderstanding (Lindblom, 1990).

Note that, unlike Lindblom (1990) and Bell et al. (2009) we do not view the lengthening process as necessarily occurring online. Given the existence of 'massive reduction', such as *probably/prolly* (Johnson, 2004) and *I don't know/ãã̃ã̃* (Bybee & Scheibman, 1999), we do not foresee that the first-mention form of a word could always be easily derived from the subsequent-mention form by a parametric adjustment of duration or any other prosodic parameter. For example, there is no parameter that could be tuned to transform *prolly* into *probably*. Thus, first- and subsequent-mention forms of a word could well be selected from a set of context-specific realizations of the word stored in memory (e.g. Bybee, 2001; Goldinger, 1998) without either one being derived from the other. In other words, while we argue that the subsequent-mention form is the production default and that the first-mention form is a departure from the default produced when the referent is mentioned in discourse for the first time (and possibly other context where clarity is desired), we do *not* propose that the first-mention form is necessarily derived from the second-mention form by an online lengthening process. The two forms could simply be alternative ways of saying the same thing triggered by different linguistic contexts (cf. the notion of a "linguistic variable" in Labov, 1969 et seq).

The proposal that first- and subsequent-mention productions are automatic productions of the words in question appropriate to the discourse context, rather than the output of online reduction and lengthening processes, is also consistent with Bard et al.'s (2000) finding that listeners were likely to provide overt cues of incomprehension when the first mentions of words were (anomalously) reduced by speakers but these cues did not induce the speaker to lengthen subsequent mentions of the misunderstood word. The first-mention form is expected to be longer by the listener (since first-mention lengthening is conventionalized). When this expectation is violated, comprehension is impaired. However, this impaired understanding, even if conveyed to the speaker, does not appear to trigger an online lengthening process.

Future work might examine the effects of repetition or first mention in circumstances when the repeatedly mentioned unit is not as likely to be automatized prior to the beginning of the experiment, including co-speech gesture and novel words in adults and low-frequency words in children. If first-mention lengthening is a learned behavior, as we suggested, children might not exhibit it in cases where adults do, especially with words they do not know well. If first-mention lengthening is not only automatic but conventional, we would hypothesize that the listener should expect it to occur, which could be tested using perception studies.

Finally, if first-mention lengthening arises out of conventionalization of a behavior that is originally listener-oriented, then co-speech gestures might be less likely to exhibit first-mention lengthening. While ubiquitous, co-speech gesture is much less conventionalized at the community level (Kendon, 2004; McNeill, 2005). On some views, it is not even communicative, as evidenced by the findings that people gesture on the phone (Bavelas et al., 2008) and congenitally blind people gesture while speaking to each other (Iverson & Goldin-Meadow, 1998). In other words, while the goal of speech is -- at least at the level of propositional content – to communicate, the goal of co-speech gesture may not be.

Conventionalized lengthening has its pre-conventionalization roots in the fact that speech is communicative, as the only reason to deviate from the low-articulatory-cost default is for the benefit of the listener (Lindblom, 1990). Given this, we might expect that the conventionalized first-mention lengthening observed in speech may not be observed in the less conventionalized, and less communicative, medium of co-speech gesture. At the same time, there is more room to automatize a novel co-speech gesture, such as one used to demonstrate the stretching of the pizza dough, over the course of the experiment. Since the gesture is novel, its production is not yet automatized at the beginning of an experiment, unlike the production of a familiar word (cf. Kapatsinski, 2010 for evidence that word production *is* automatized). As a result, gestures (particularly, mimetic gestures such as the one depicting the stretching of a pizza dough) may reduce across stories (regardless of listener) even though the words they co-occur with (their 'lexical affiliates') do not. Preliminary results from Vajrabhaya & Pederson (2013, 2014) support this conclusion.

Conclusion

With the exception of Fowler et al. (1997), past literature has examined the effect of repetition on word duration as an online process of reduction or, less commonly, lengthening (Bell et al., 2009). Perhaps, it is critical that we, as researchers, rethink and reexamine production more broadly. Speakers' behavior is often motivated by convention, including discourse convention. These conventions arise from the online processing pressures resulting from difficulties with lexical access, utterance comprehension and speech planning (e.g. Givón, 1979). However, they codify some pressures over others, allowing the speaker to track fewer variables during production. Conventionalizing a processing pressure frees the speaker from the processing nightmare of online adjustment of production plans based on resource-demanding real-time modeling of the listener. However, it also means that the speaker's production will not be exquisitely sensitive to the listener's needs.

Thus, we document that speakers' production of words that are mentioned for the first time within a coherent stretch of discourse ('story') are longer than subsequent mentions of the same word within that story or preceding mentions of the word within the preceding story. This first-mention lengthening occurs even if the word has previously mentioned for the same listener and is therefore easily accessible to both speaker and listener. We argue that the pattern of first-mention lengthening is automatic and conventionalized. Since the listener normally needs the speaker to be extra clear when a word is mentioned for the first time in a discourse, the extra clarity at first mention becomes conventionalized as how one tells a story. A native speaker of English or Thai then has to learn to provide the extra clarity as part of acquiring their native language, to the point of doing so automatically, even in the rare situations when the listener does not in fact need it. There is no harm in needlessly lengthening a word mentioned for the first time while not lengthening when lengthening is needed can result in misunderstanding (as documented by Bard et al., 2000). Thus, the speaker has motivation to generalize lengthening beyond the contexts where it is strictly speaking necessary, to the point of lengthening automatically whenever a word is mentioned for the first time within a discourse.

Acknowledgements

This project was partly funded by the Southeast Asian Studies Award from the University of Oregon to the first author and a start-up grant from the University of Oregon to the second author. We would like to thank Eric Pederson and Alexia Galati their insightful comments and useful feedback. We would also like to thank Tyler Kendall for providing recording space in the Language and Variation and Computation Lab.

References

Abramson, A. (2002). The stability of distinctive vowel length in Thai. In K. Tingsabadh & A. Abramson. (Eds.), *Essays in Tai linguistics* (pp.13-26). Bangkok: Chulalongkorn University Press.

Anderson, A. H., & Howard, B. (2002). Referential form and word duration in video-mediated and face-to-face dialogues. In J. Bos, M. E. Foster, C. Matheson (Eds.), *Proceedings of EDILOG 2002*. Edinburgh: Cognitive Science Centre, University of Edinburgh.

Aylett, M., & 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*, 31-56.

Baker, R. E. & Bradlow, A. R. (2009). Variability in word duration as a function of probability, speech style, and prosody. *Language and Speech, 52,* 391-413.

Bard, E. C., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42*, 1-22.

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278.

Bates, D. (2013). *Linear mixed model implementation in lme4*. University of Wisconsin, Madison, WI.

Bavelas, J. & Healing, S. (2013). Reconciling the effect of mutual visibility on gesturing: A review. *Gesture, 13*, 63-92.

Bavelas, J., Gerwing, J., Sutton, C. & Prevost, D. (2008). Gesturing on the telephone: independent effects of dialogue and visibility. *Journal of Memory and Language, 58*, 495-520.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language, 60*, 92-111.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer (version 5.1.25)[Software]. Available from: http://www.praat.org/.

Broadbent, D. E. (1967). Word-frequency effect and response bias. *Psychological Review, 74*, 1-15.

Butler, B. C. (2012) *A semantic map approach to English articles (a, the, and Ø).* (Unpublished Doctoral Dissertation). University of Oregon, Eugene, OR.

Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language, 82*, 711-733.

Bybee, J. (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change, 14*, 261-290.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J., & Scheibman, J. (1999). The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics*, *37*, 575–596.

Byrd, D. M. (1994). Articulatory timing in English consonant sequences. Ph.D. dissertation, UCLA, distributed as *UCLA Working Papers in Phonetics*, *86,* 1-196.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. Li (Ed.), *Subject and topic* (pp.25–55). New York: Academic Press.

Clark, H. H. & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. Joshi, B. H. Weber & I. A. Sag (Eds.), *Elements of discourse understanding* (pp.414-460). Cambridge: Cambridge University Press.

DeLancey, S. (2011). Grammaticalization and syntax – A functional view. In H. Narrog & B. Heine (Eds.), *The Oxford handbook of grammaticalization* (pp.365-377). Oxford: Oxford University Press.

Erker, D., & Guy, G. R. (2012). The role of lexical frequency in syntactic variability: Variable subject personal pronoun expression in Spanish. *Language*, *88*, 526-557.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 680-698.

Fougeron, C., & P. Keating. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America, 101*, 3728-3740.

Fowler, C. (1988). Differential shortening of repeated context words produced in various communicative contexts. *Language and Speech, 31*, 307-319.

Fowler, C., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 62*, 35-51.

Fowler, C. A., Levy, E. T., & Brown, J. M. (1997). Reductions of spoken words in certain discourse contexts. *Journal of Memory and Language, 37*, 24-40.

Fox Tree, J. E. (2010). Discourse markers across speakers and settings. *Language and Linguistics Compass*, *4*, 269-281.

Frank, A., & Jaeger, F. (2008). Speaking rationally: Uniform information density as an optimal strategy for language production. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 939-944). Austin, TX: Cognitive Science Society.

Freeman, V. (2014). Hyperarticulation as a signal of stance. *Journal of Phonetics, 45*, 1-11.

Gahl, S. (2008). "Thyme" and "time" are not homophones. The effect of lemma frequency on word durations in spontaneous speech. *Language, 84*, 474-496.

Galati, A., & Brennan, S. E. (2014). Speakers adapt gestures to addressees' knowledge: implications for models of co-speech gesture. *Language, Cognition, and Neuroscience*, *29*, 435-451.

Galati, A. & Brennan, S. E. (2010). Attenuating repeated information: For the speaker, or for the addressee? *Journal of Memory and Language, 62*, 35-51.

Givón, T. (1979). *On understanding grammar*. New York: Academic Press.

Goldiamond, I & Hawkins W.F. (1958). Vexierversuch: The logarithmic relationship between word-frequency and recognition obtained in the absence of stimulus words. *Journal of Experimental Psychology*, *56*, 457–463.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251-279.

Haberlandt, K., Berian, C., & Sandson, J. (1980). The episode schema in story processing. *Journal of Verbal Learning and Verbal Behavior, 19*, 635-650.

Howes, D. H. & Solomon, R. L. (1951). Visual duration threshold as a function of word probability. *Journal of Experimental Psychology, 41*, 401-410.

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground. *Cognition, 59*, 91-117.

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture when they speak. *Nature*, *396,* 228.

Jaeger, F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, *61*, 23-62.

Johnson, K. (2004). Massive reduction in conversational American English. In K. Yoneyama & K. Maekawa (Eds.), *Spontaneous Speech: Data and Analysis. Proceedings of the 1st session of the 10th International Symposium* (pp.29-54). Tokyo: The National International Institute for Japanese Language.

Katz, J., & Selkirk, E. (2011). Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*, *87*, 771-816.

Kapatsinski, V. (2010). Frequency of use leads to automaticity of production: Evidence from repair in conversation. *Language and Speech, 53*, 71-105.

Keating, P., Cho, T., Fougeron, C. & Hsu, C-S. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden & R. Temple (Eds.). *Papers in Laboratory Phonology VI* (pp.143-161). Cambridge: Cambridge University Press.

Keating, P., Wright, R., & Zhang, J. (1999). Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics, 97*, 157-173.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kuhlen, A. K. & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin and Review, 20*, 54-72.

Labov, W. (1969). Contraction, deletion, and inherent variability in the English copula. *Language, 45,* 715-762.

Ladd, R. (1996). *Intonational phonology*. New York: Cambridge University Press.

Lam, T. Q. & Watson, D. G. (2010). Repetition is easy: Why repeated referents have reduced prominence. *Memory and Cognition, 38*, 1137-1146.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of H&H theory. In W. Hardcastle & A. Marchal. (Eds), *Speech production and speech modeling* (pp.403-439). Dordrecht: Kluwer.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95*, 492-527.

McNeill, D. (2005). *Gesture and thought*. Chicago: University of Chicago Press.

Mol, E. M. M., Krahmer, E. J., Maes, A., & Swerts, M. G. J. (2009). The communicative import of gesture:Evidence from a comparative analysis of human-human and human-machine interactions. *Gesture, 9*, 98-127.

Myers, Scott & Jaye Padgett (to appear). Domain generalization in artificial language learning. *Phonology*.

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*, 357-395.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology, 4*, 272-281.

Patterson, D. J., & Connine, C. M. (2001). A corpus analysis of variant frequency in American English flap production. *Phonetica, 58*, 254–275.

Pickering, M. J., & Garrod, S. (2004). Towards a mechanistic psychology of dialog. *Behavioral and Brain Sciences, 27*, 169-190.

Pierrehumbert, J. (2001). Why phonological constraints are so coarse-grained. *Language and Cognitive Processes*, *16*, 691-698.

Pierrehumbert, J., & Talkin, D. (1991) Lenition of /h/ and glottal stop. In G. Docherty, & D.R. Ladd (Eds.), *Papers in Laboratory Phonology II* (pp.90-117). Cambridge, UK: Cambridge University Press.

Pluymaekers, M., Ernestus, M., Baayen, R. H. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica, 62*, 146-159.

R Development Core Team. (2009). R: A language and environment for statistical Computing. Vienna: R Foundation for Statistical Computing. Available from: http://www.R-project.org.

Schuppler, B., Ernestus, M., Scharenborg, O., & Boves, L. (2011). Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions. *Journal of Phonetics, 39*, 96–109.

Seibel, R. (1963). Discrimination reaction time for a 1,023 alternative task. *Journal of Experimental Psychology, 66*, 215-226.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Biases in judgments reveal some heuristics of thinking under uncertainty. *Science, 185*, 1124-1131.

den Uyl, Martyn, & Herre van Oostendorp. (1980). The use of scripts in text comprehension. *Poetics*, *9*, 275-294.

Vajrabhaya, P., & Kapatsinski, V. (2011). There is more to the story: First-mention lengthening in Thai interactive discourse. In W. S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences,* 2050-2053. Hong Kong.

Vajrabhaya, P., & Pederson, E. (2014). When anyone's listening: Automatization and gesture reduction. Paper presented at ISGS6, San Diego.

Vajrabhaya, P., & Pederson, E. (2013). Repetition vs. listener accommodation: A case study of co-speech gesture in retellings. Paper presented at the NWAV42, Pittsburgh.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Oxford: Addison-Wesley Press.