Journal of Experimental Psychology:
Learning, Memory, and Cognition

ISSN: 0278-7393

2020, Vol. 46, No. 11, 2163–2178
http://dx.doi.org/10.1037/xlm0000939

The Scope of Audience Design in Child-Directed Speech: Parents' Tailoring of Word Lengths for Adult Versus Child Listeners

Nicholas Tippenhauer
Vanderbilt University

Eva R. Fourakis
Princeton University

Duane G. Watson
Vanderbilt University

Casey Lew-Williams
Princeton University

When communicating with other people, adults reduce or lengthen words based on their predictability, frequency, and discourse status. But younger listeners have less experience than older listeners in processing speech variation across time. In 2 experiments, we tested whether English-speaking parents reduce word durations differently across utterances in child-directed speech (CDS) versus adult-directed speech (ADS). In a child-friendly game with an array of objects and destinations, adult participants (N = 48) read instructions to an experimenter (adult-directed) and then to their own 2- to 3-year-old children (child-directed). In Experiment 1, speakers produced sentences containing high-frequency target nouns, and in Experiment 2, they produced sentences containing low-frequency target nouns. In both CDS and ADS in both experiments, speakers reduced repeated mentions of target nouns across successive utterances. However, speakers reduced less in CDS than in ADS, and low-frequency nouns in CDS were overall longer than low-frequency nouns in ADS. Together, the results suggest that repetition reduction may be beyond speaker control, but that speakers still engage in audience design when producing words for relatively inexperienced listeners. We conclude that language production involves nested audience-driven and speaker-driven processes.

Keywords: child-directed speech, language production, word duration, prosody, discourse status

Humans are capable of sharing information efficiently and dynamically during natural communication, and many factors affect their success in doing so. For example, shared information in common ground (e.g., Stalnaker, 1973) and interlocutors' perspectives (e.g., Brown-Schmidt & Heller, 2018) can influence exchanges between two people. Successful communication is also constrained by subtle temporal aspects of speech. For example, when conveying messages to others, we reduce and lengthen words differently depending on their predictability, frequency, and discourse status (e.g., Aylett & Turk, 2004; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Fowler & Housum, 1987; Jurafsky, Bell, Gregory, & Raymond, 2001; Lam & Watson, 2010). We also make adjustments to speech depending on the identity of our listeners, spanning native adult speakers, non-native adult speakers, children, and even animals (Schwab & Lew-Williams, 2016a; Uther, Knoll, & Burnham, 2007; Xu, Burnham, Kitamura, & Vollmer-Conna, 2013). Importantly, these factors interact with one another, and speakers must navigate repetition of information with the needs of their listeners (e.g., Bortfeld & Morgan, 2010). Researchers who study language production debate whether such audience design is dictated by the needs of speakers or by the needs of their audiences (e.g., Arnold, Kahn, & Pancani, 2012; Aylett & Turk, 2004; Bard et al., 2000; Jaeger, 2010). Here, in order to examine this debate, we compared parents' word use across utterances in the established context of adult–adult communication and the underexamined context of parent–child communication. In doing so, this investigation furthers what is known about language production and about audience design in child-directed speech.

A word's discourse status, that is, whether it is new or given in the conversation, plays an important role in shaping speakers' productions (e.g., Aylett & Turk, 2004; Chafe, 1987; Prince, 1981). Newly referenced words tend to be longer and louder, whereas given references tend to be shorter and quieter (e.g., Bard et al., 2000; Bell et al., 2009; Fowler & Housum, 1987; see

This article was published Online First July 23, 2020.

Nicholas Tippenhauer, Department of Psychology and Human Development, Vanderbilt University; Eva R. Fourakis, Department of Psychology, Princeton University; Duane G. Watson, Department of Psychology and Human Development, Vanderbilt University; Casey Lew-Williams, Department of Psychology, Princeton University.

This project was funded in part by grants from the National Institute of Child Health and Human Development (R01HD095912, R03HD079779) to Casey Lew-Williams and a grant from the National Science Foundation (NSF-BCS 15570797) to Duane G. Watson. We thank Maritza Gomez for help with data collection. We would also like to thank participating families for taking part in our research.

Correspondence concerning this article should be addressed to Nicholas Tippenhauer, Department of Psychology and Human Development, Vanderbilt University, 230 Appleton Place PMB 552, Nashville, TN 37203. E-mail: n.tippenhauer@vanderbilt.edu

Wagner & Watson, 2010 for a review). This given/new distinction has left the field with a puzzle: Why is word duration influenced by discourse status? There is general agreement that the link between discourse status and word duration facilitates communication, but there is disagreement as to who benefits from this link. And while it is clear that repeating a word causes repetition reduction, it is unclear why exactly this is the case. By exploring duration differences based on the discourse status in both child-directed speech (CDS) and adult-directed speech (ADS), we aimed to better understand which of two predominant explanations is correct.

One explanation is that changes in word duration reflect speakers' production system—specifically their planning processes (e.g., Arnold & Watson, 2015; Bell et al., 2009; Kahn & Arnold, 2012, 2015). The central idea is that because speech is planned incrementally, speakers adjust word durations with planning time in mind; words that are difficult to produce are lengthened, and words that are more easily produced are shortened. Evidence for these claims comes from work demonstrating that frequency and predictability correlate with word duration (Bell et al., 2009), presumably because it is easier to produce something that is commonly mentioned or likely to be mentioned. In addition, there is evidence that priming at any level of the production process (i.e., conceptual, syntactic, lexical, phonological, or articulatory levels) leads to reduction (Kahn & Arnold, 2012, 2015). Both Bell et al. (2009) and Kahn and Arnold (2012, 2015) argue that these effects are due to ease of retrieval for the speaker.

A second possibility is that speakers modulate word duration to optimize language comprehension for their listeners (e.g., Aylett & Turk, 2004; Jaeger, 2010; Lieberman, 1963). Here, the main idea is that repeated, more predictable words are reduced because they are easy to comprehend, whereas new, less predictable words are lengthened because they are more difficult to process (Lieberman, 1963). This idea has been formalized in information theoretic accounts, which argue that speakers attempt to maintain consistent information density throughout an utterance (e.g., Aylett & Turk, 2004; Jaeger, 2010; Pate & Goldwater, 2015). By lengthening high-information words, such as words that are unpredictable or new, and shortening low-information words, such as words that are predictable or given, speakers can transmit a constant amount of information over time. In short, these accounts argue that speech modulation is an act of audience design, a linguistic process intended to benefit listeners' comprehension.

Disentangling these competing accounts is difficult. Linguistic material that is easy for a speaker to produce also tends to be easy for a listener to comprehend. One strategy for decoupling these possibilities is examining conversations in which one of the interlocutors is less knowledgeable, experienced, or competent. If speakers tailor their speech for this less knowledgeable audience, even if it comes at some cost to the speaker, it would suggest that audience design plays a key role in articulation. Adult speakers do alter certain aspects of speech for less knowledgeable listeners, such as non-native speakers (e.g., Scarborough, Dmitrieva, Hall-Lew, Zhao, & Brenier, 2007; Smith, 2007; Uther et al., 2007) and children through CDS (see Soderstrom, 2007, for a review). However, only one study has examined how speakers modify word lengths for more versus less knowledgeable listeners (Fisher & Tokura, 1995).

CDS is characterized by a number of unique features that may support children's learning and understanding of language, including higher pitch, more variable pitch, exaggerated vowels, more repetition, shorter utterances, shifted vocal timbre, and musical intonation (e.g., Cristia, 2013; Piazza, Iordan, & Lew-Williams, 2017; Soderstrom, 2007; Schwab & Lew-Williams, 2016a). Perhaps because of these features, CDS seems to capture the attention and interest of children from an early age (e.g., Cooper & Aslin, 1990; Fernald, 1985; Pegg, Werker, & McLeod, 1992; Werker & McLeod, 1989; Werker, Pegg, & McLeod, 1994). Importantly, features of CDS facilitate language learning and development. For example, caregiver speech clarity is related to speech discrimination (Liu, Kuhl, & Tsao, 2003), and repetition of labels across successive utterances—a hallmark of CDS—facilitates better encoding of new words (Schwab & Lew-Williams, 2016b). In addition, infants are better at segmenting word-like units when presented with CDS than when presented with ADS (Thiessen, Hill, & Saffran, 2005; although see Bard & Anderson, 1983 who found that CDS may be less intelligible than ADS). Regardless, given the boost to learning and comprehension that CDS affords children, it is highly likely that speakers produce CDS for their audience (Schwab, Rowe, Cabrera, & Lew-Williams, 2018). In other words, CDS appears to be an extreme version of audience design intended to promote listener comprehension.

Because CDS is clearly designed to optimize perception, communication, and learning for a particular audience, it is an ideal testing ground for probing the limits of audience design. Understanding how word reduction and lengthening function in CDS versus ADS may shed light on whether these features of communication are speaker- or listener-centered. If reduction is speaker-centered, one might expect reduction even in CDS, because production constraints would determine word duration independent of the audience. If reduction is listener-centered, one might expect reduction to be attenuated or even eliminated in CDS, because relatively less knowledgeable listeners need clearer speech (compared to more knowledgeable listeners), independent of the predictability or discourse status of words.

The question of whether speakers reduce repeated information in CDS was investigated over 20 years ago by Fisher and Tokura (1995). In their study, caregivers were asked to spontaneously describe scenes from a puppet show to either an adult experimenter or to their own 14-month-old child. Repeated targets were those that occurred in two successive utterances, and new targets were those that differed across successive utterances. Fisher and Tokura (1995) found that words tended to be longer in CDS compared to ADS, but reduction of repeated, or given references was evident in both types of speech. In addition, they found more reduction in CDS than ADS. However, they acknowledge that this difference was likely due to caregivers' greater tendency in CDS than in ADS to place target words at the ends of sentences, which tend to be lengthened in general (Ratner, 1986; see Wagner & Watson, 2010 for a review). That is, simply because utterance-final words tend to be longer, more absolute reduction was possible in CDS than ADS—independent of audience design (e.g., Fisher & Tokura, 1995). Based on these findings, Fisher and Tokura (1995) argued that while speakers can actively control broad changes to speech, like choosing to engage in CDS, they cannot control finer-grained changes, such as reduction of repeated words. A related study conducted by Bortfeld and Morgan (2010) found similar evidence:

In CDS, adults alternatively reduced and lengthened target word across successive utterances. Thus, there may be a complicated interplay between audience design and automatic processes in speech production.

In light of recent debates about the nature of reduction (Aylett & Turk, 2004; Jaeger, 2010), we designed two experiments to evaluate parents' reduction in CDS and ADS. We did this for two reasons. The primary reason was to evaluate whether speakers, as in Fisher and Tokura (1995), actually do reduce repeated words more in CDS than in ADS. They argued that speakers have broad control over how an utterance is articulated, such as deciding whether to engage in CDS or not, but that audience design is not a factor in fine-grained, word-specific production processes. Thus, in investigating this claim, we aimed to better understand the scope of audience design. A second, related reason was to build better stimulus control into our design, because Fisher and Tokura (1995) used a free-form procedure that, though natural, introduced potential production artifacts. Speakers were instructed to describe scenes they had witnessed, which could have resulted in variable amounts of information separating repeated references, and as discussed above, this generated sentence-position inequality between CDS and ADS. Indeed, Fisher and Tokura (1995) acknowledged that differences in the amount of reduction between ADS and CDS may have been due to differences in participants' unconstrained speech behaviors. In addition, simply restating information, as in the case of caregivers describing scenes to their infants, may not engage certain elements of audience design. Recent work showed that speakers are less sensitive to a listeners' perspectives when the speakers are simply providing information versus making a request (Yoon, Koh, & Brown-Schmidt, 2012).

Based on prior research, we generated two main hypotheses. First, we hypothesized that speakers automatically engage in fine-tuned durational changes including repetition reduction regardless of their audience. But second, we hypothesized that speakers make broad changes to their speech when addressing different audiences, such as the use of CDS with young children. We conducted two experiments to test these hypotheses. In both experiments, using a child-friendly game with an array of objects and destinations, participants read instructions to an adult and then to their own child. Instructions included referents that were either repeated across successive utterances or newly mentioned across utterance pairs (see Figure 1). In Experiment 1, participants read instructions that included high-frequency target nouns (e.g., bird). In Experiment 2, to further probe the limits of adults' audience design, they read instructions that included low-frequency target nouns (e.g., brooch). If adults modulate discourse reduction to help their child audience, they should show less reduction of repeated words when speaking to children than to adults, and, in turn, they should show less reduction for low-frequency words than for high-frequency words. If adults can make broad changes to speech but are nonetheless constrained by production-centered processes (Fisher & Tokura, 1995), they should make general audience-based changes to speech but should not have control over how much they reduce repeated words.

Experiment 1

Method

Participants. Participants were 24 English-speaking parents (23 female) and their 24 young children (13 female, M = 38.1 month, range = 32 to 48 months). All children were exposed to English at least 80% of the time (M = 97.7%). No child had a history of hearing problems or pervasive developmental delays. Seven additional parent–child dyads were tested but excluded from analyses due to children's inattentiveness (i.e., not sitting through at least half of the utterances, n = 4) or parents' failure to follow directions (n = 3). This study was approved by the Insti-

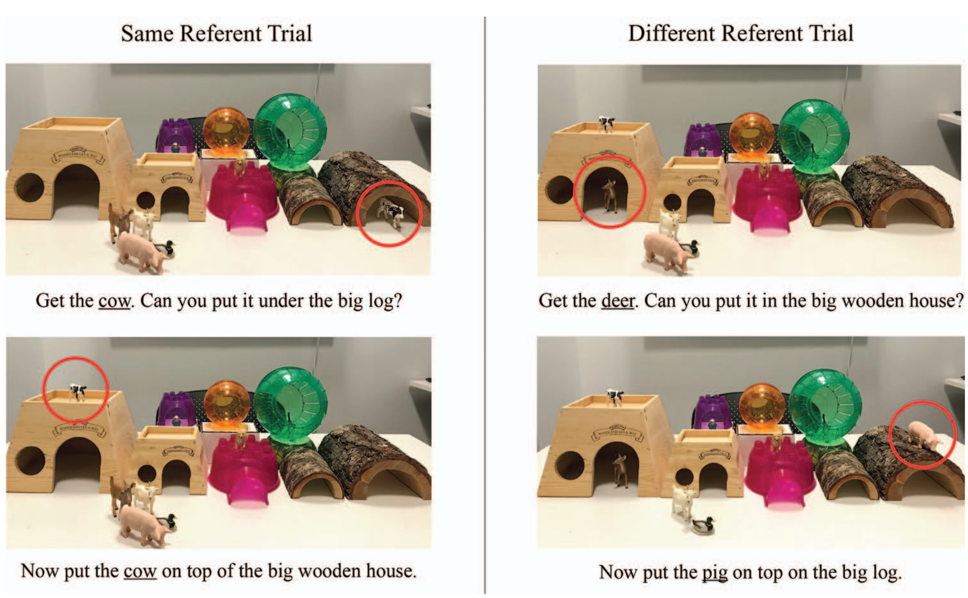

Figure 1. Example same referent and different referent trials and the array of target animals and locations used in Experiment 1. See the online article for the color version of this figure.

tutional Review Board at Princeton University (IRB #7117, "Language Learning").

Stimuli and design. Each parent's task was to instruct an adult experimenter and then their child to place small plastic animals into an array of objects on a table. The eight target animals were presumed to be high-frequency labels in input to 3-year-old children: bird, cat, cow, deer, dog, duck, goat, and pig. The array consisted of eight possible destinations for the animals: a little wooden house, big wooden house, little hamster ball, large hamster ball, little log, big log, purple castle, and pink castle. A sequence of 32 utterances was arranged into 16 pairs of instructions (see Appendix A). Eight pairs of instructions were part of Same Referent trials, in which parents repeated the same animal name in consecutive instruction, for example, "Get the *cow*. Can you put it under the big log? Now put the *cow* in the purple castle." The other eight pairs were part of Different Referent trials, in which parents produced a different animal name in consecutive instructions, for example, "Get the *cow*. Can you put it in the little ball? Now put the *deer* in the big wooden house." Targets appeared in sentence-final positions in first utterances and in sentence-medial positions in second utterances.

Each animal name appeared the same number of times on same referent and different referent trials, and the same number of times in the position of first and second mention within a pair of instructions (see Appendix A). For example, *deer* was used in both first and second mention on one same referent trial, first mention on one different referent trial, and second mention on a separate different referent trial. The 16 instruction pairs alternated between same referent and different referent trials. Instructions were pseudorandomly ordered such that no animal name appeared in any sequence of two instructions, and no sequential instructions ever referred to the same location.

Procedure. Participants viewed an array of eight animal figurines and eight locations arranged on a table (see Figure 1). They were also given a list of 32 single-sentence instructions (see Appendix A). Parents were told that they would be reading these simple instructions to their child in order to see how children respond to their parents at various ages. Parents were instructed to avoid straying from the list of sentences and to avoid repeating any sentences, even if their child (a) did not respond, (b) picked up an incorrect object, or (c) placed an object in an incorrect location.

Parents first read the list of instructions to an adult experimenter (without the child present) to place animals in specific locations. This session, which lasted 3.2 min on average, was described as practice for the subsequent task with children. After reuniting, parents instructed their children using an identical sequence of sentences, which lasted 9.8 min on average. Parents' speech was recorded using a lapel microphone, and parents were seated next to the experimenter or the child during each game.

Coding. Coders blind to condition segmented target words using spectrograms and waveforms using Praat (Boersma & Weenink, 2017). Coders were presented with whole utterances for each mention in isolation and instructed to segment only target words and to exclude sounds from the words surrounding the target. The coders used standardized guidelines for identifying different phonemes in Praat. In every case, the beginning of a target's initial consonant indicated the onset of that target's duration. Likewise, the end of a target's final consonant indicated the offset of that target's duration. Example spectrograms were provided to help coders identify consonants within the speech stream. Coders were told to visually inspect the spectrogram for the beginning (and end) of energy associated with the target. They were also told to use the audio to help guide their judgments. This segmentation yielded target word durations for each utterance.

Results

The main questions addressed in this experiment were whether or not adults would (a) reduce the lengths of repeated words in CDS and ADS and (b) show less reduction of repeated words in CDS than in ADS. Within each pair of utterances, the key dependent variable was the duration of the target noun in the first versus second utterance. Only first and second mentions of words in each pair were included for analysis. Any additional repetitions of target nouns were not included in these analyses, as they represented a departure from the verbal instructions. Across all participants, the total number of target words was 1,536. We excluded 140 targets (9%) due to caregivers' disfluencies or use of incorrect labels. Durations were scaled, mean-centered, and log-transformed to control for positive skew. Nontransformed results for Experiment 1 are depicted in Figure 2.

We constructed maximal mixed-effects models with random slopes and intercepts by items and participant using the R (Version 3.4.3) package lmerTest (Version 3.0–1).We included audience (ADS or CDS), utterance position (first or second), and target word status (same referent or different referent) as factors. However, the maximal model did not converge for these data. We constructed additional models and systematically removed interaction terms from our random effects in order to generate simpler models. We used the ANOVA function in R to compare different models. The model with the smallest AIC was selected. This model included random slopes for a target word status by utterance position interaction and a separate random slope for audience by both participant and target label. It is important to note that reported coefficients do not map directly onto millisecond differences because durations were log-transformed. Output for all models can be found in Appendix B.

We first constructed a model for all of the duration data across both ADS and CDS. Using this model, we found evidence that target nouns in first utterance positions ($M = 0.46$ s, $SE = 0.02$ s) were longer overall than in second utterance positions ($M = 0.40$ s, $SE = 0.03$ s), $t = 7.94$, $p < .0001$. We did not find evidence that targets in CDS ($M = 0.46$ s, $SE = 0.03$ s) were reliably longer than in ADS ($M = 0.40$ s, $SE = 0.02$ s), $t = 1.08$, $p = .29$. However, we found a significant two-way interaction between utterance position and target word status, $t = 4.16$, $p < .001$. In addition, we found a significant three-way interaction between audience, utterance position, and target word status, $t = 2.38$, $p = .02$. We constructed additional models to explore these interactions.[1]

We constructed models for ADS and CDS trials separately in order to determine whether or not repetition reduction occurred in both types of speech. For ADS, we found that target nouns in

[1] We also constructed a model that included Global Mention (i.e., whether it was a target's first, second, third, or fourth appearance within the instructions overall) as a fixed effect. In this analysis, the main effect of global mention was not significant. Including this effect did not meaningfully alter the results, so we do not report it here.

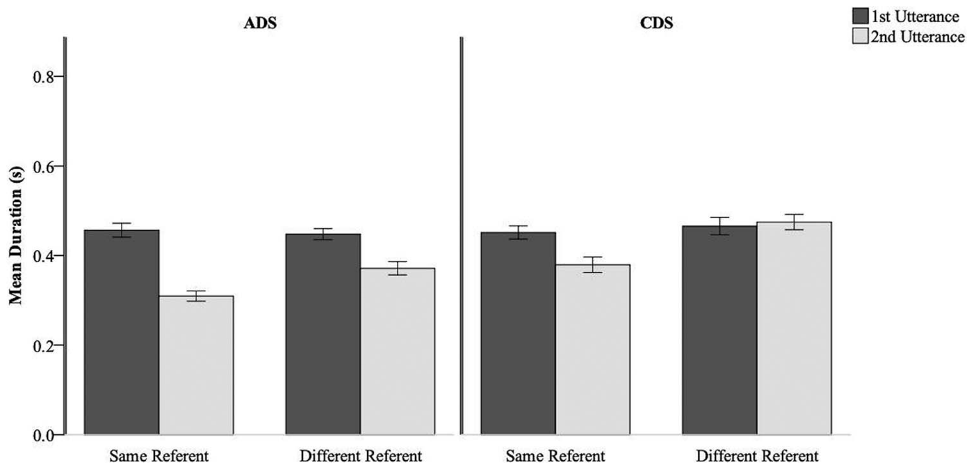

Figure 2. Target noun durations, in seconds, for first and second utterances by audience (ADS or CDS) and target word status (same referent or different referent). ADS = adult-directed speech; CDS = child-directed speech.

second utterance positions (M = 0.34 s, SE = 0.02 s) were shorter on average than targets that occurred in first utterance positions (M = 0.45 s, SE = 0.02 s), $t = -6.78$, $p < .0001$. However, an interaction between utterance position and target word status revealed that there was more reduction of second utterance targets in same-referent pairs than second utterance targets in different-referent pairs, $t = 3.81$, $p = .002$. A planned comparison revealed that second mentions of targets on same referent trials (M = 0.31 s, SE = 0.01 s) were shorter than first mentions (M = 0.46 s, SE = 0.02 s), $t = -6.93$, $p < .0001$. An additional planned comparison revealed that second mentions of target nouns on different referent trials (M = 0.37 s, SE = 0.02 s) were also shorter than first mentions (M = 0.45 s, SE = 0.01 s), $t = -3.34$, $p = .004$. Thus, for ADS, speakers reduced the names of referents regardless of whether they were the same or different referent as in the previous utterance, but they tended to reduce same-referent words more. We return to this in the Discussion.

We next constructed a model for CDS word durations. This model revealed that second mentions of target nouns (M = 0.45 s, SE = 0.02 s) were reliably shorter than first mentions of target nouns (M = 0.48 s, SE = 0.02 s), $t = -5.26$, $p < .0001$. We also found an Utterance Position by Target Word Status interaction, suggesting that reduction between first and second utterance positions for same referent pairs was greater than reduction between first and second utterance positions for different referent pairs, t = 5.51, $p < .0001$. Planned comparisons revealed that second mentions (M = 0.38 s, SE = 0.03 s) of nouns on same referent trials were reliably shorter than first mentions (M = 0.48 s, SE = 0.03 s), $t = -5.40$, $p < .0001$. However, first mentions of nouns on same referent trials (M = 0.49 s, SE = 0.03 s) were statistically no different in duration than second mentions (M = 0.51 s, SE = 0.03 s), $t = 0.38$, $p = .71$. In CDS, speakers only reduced words referring to the same target as in the previous utterance.

Finally, we constructed a model comparing same referent trials across ADS and CDS. We found a significant interaction between audience and utterance position, $t = 2.39$, $p = .04$, indicating that there was less reduction of same referent words in CDS than in ADS. Parents did not reduce the durations of repeated words as much when speaking to their children.

Discussion

We found evidence in Experiment 1 that adults reduced repeated words regardless of their audience. But critically, they reduced repeated references less when talking to children than when talking to adults. These results both confirm and extend previous work demonstrating the effects of discourse status on word duration (e.g., Bard et al., 2000; Bell et al., 2009; Fowler & Housum, 1987; see Wagner & Watson, 2010). However, our results conflict with previous investigations of reduction in CDS (e.g., Fisher & Tokura, 1995), which found more reduction in CDS compared to ADS.

Our experiment, in tandem with Fisher and Tokura (1995), could suggest that repetition reduction is an automated process. Adult speakers in Experiment 1 reduced repeated mentions of words despite the fact that children may have benefited from hearing nonreduced versions of target words. This could be because repetition reduction is an inherent feature of speakers' production systems that does not—and perhaps cannot—consider audience needs during real-time communication (e.g., Arnold & Watson, 2015; Bell et al., 2009; Kahn & Arnold, 2012, 2015). This is in line with Fisher and Tokura's (1995) prediction that speakers may not have control over fine-grained production processes.

However, we found evidence that speakers treated adult and child audiences differently. Speakers reduced repeated references less in CDS than in ADS (cf. Fisher & Tokura, 1995, who found that reduction was greater for CDS than ADS). The difference between our results and Fisher and Tokura (1995) is possibly due to differences in experimental methods; in Fisher and Tokura (1995), speakers were free to describe events without constraint and often placed targets at the end of utterances in CDS but not ADS. Previous research has demonstrated that utterance-final words often receive lengthening irrespective of other factors (e.g., Ratner, 1986; Fisher & Tokura, 1995; see Wagner & Watson, 2010). Our more controlled design allowed us to compare sentences where target nouns occurred in the same position across both ADS and CDS, and we find different results.

It is worth noting, however, that targets in our study differed in sentence placement across first and second utterances. Targets occurred at the end of first utterances (e.g., "Can you get the *dog*?") and in the middle of second utterances (e.g., "Now, put the *dog* in the little wooden house."). This difference may explain why we see reduction in the same referent condition in ADS. While it is not optimal that targets occurred in different syntactic positions across utterances, this design does not undermine our ability to detect repetition reduction, and it does not prevent us from drawing conclusions about relative reduction between ADS and CDS. This design feature was necessary in order to discourage pronoun use across utterances; there is existing evidence that speakers prefer using pronouns to refer to a repeated noun if it is used in the same syntactic position (e.g., Chambers & Smyth, 1998; Smyth, 1994). The fact that this pattern of reduction occurred only in ADS and not in CDS provides evidence that speakers can tailor their speech to accommodate their audience's needs.

The fact that we found different patterns of reduction in ADS and CDS challenges the claim made by Fisher and Tokura (1995) that speakers do not have control over automatic production processes (e.g., repetition reduction). A revision to Fisher and Tokura's (1995) hypothesis is warranted based on the current data. Namely, speakers may not have active control over whether or not they engage in word-specific, fine-grained reduction, but they are able to vary the extent to which they reduce for children versus adults. Essentially, we found evidence for nested processes that include both automatic constraints and audience design.

It is important to note that we did not find any target duration differences overall between ADS and CDS in Experiment 1—in contrast to a reliable duration difference in Fisher and Tokura (1995). Two differences between the present study and Fisher and Tokura (1995) may explain this contrast. First, our 2- and 3-year-old participants were older than the 14-month-old infants tested in Fisher and Tokura (1995). We opted to test older children because there is existing evidence that speakers engage in more audience design if their listener completes an action (Yoon et al., 2012), and we aimed to simulate a dynamic, shared interaction between adults and young children. Moreover, there is evidence that adults engage in audience design for young children of various ages (Schwab & Lew-Williams, 2016a), and there was no evidence-based reason to believe that caregivers' word durations would differ significantly across the first years of life. Second, target words in Experiment 1 were high frequency in the environments of young children, whereas targets in Fisher and Tokura's (1995) study were unknown to infant listeners. Perhaps adults only engage in word-duration reduction in CDS when producing labels that are likely to be familiar to children, because they recognize that these are contexts in which children may not need scaffolding for successful comprehension. We tested this possibility in Experiment 2 by having participants read instructions with low-frequency nouns that referred to less common objects, rather than the high-frequency targets used in Experiment 1. This also allowed us to replicate the general findings from Experiment 1.

Experiment 2

In Experiment 2, we tested whether repetition reduction occurred when parents produced labels for relatively uncommon objects that were less likely to be familiar to their children. The use of low-frequency target nouns in Experiment 2 (vs. high-frequency nouns in Experiment 1) provided a more complete evaluation of parents' repetition reduction in CDS and ADS, as parents may be less likely to reduce words that are less familiar to their children.

Method

Participants. Participants were 24 English-speaking parents (21 female) and their 24 young children (14 female, $M = 42.0$ months, range = 33 to 48 months). Participants who completed Experiment 1 did not participate in Experiment 2. All children were exposed to English at least 80% of the time ($M = 96.1\%$). No child had a history of hearing problems or pervasive developmental delays. Five additional parent–child dyads were tested but excluded from analyses due to inattentiveness ($n = 3$) or experimenter error ($n = 2$).

Stimuli and design. The target objects used in Experiment 2 were eight uncommon objects (as opposed to common animals in Experiment 1) that children were unlikely to produce on a regular basis, or at all: brooch, felt, globe, gnome, spade, spool, thimble, and twine. These words were selected based on their very low occurrence in North American English corpora on the Child Language Data Exchange System (range 0–39 occurrences out of 4,271,640 total words; MacWhinney, 2000). All other stimuli were identical to Experiment 1.

Procedure. The procedure was nearly identical to Study 1, except that we told parents they could point to objects, as children were unlikely to know which label referred to which object.

Coding. The lengths of first- and second-mention nouns were determined using the same procedures as in Experiment 1.

Results

Again, our key dependent variable was the duration of target nouns for each pair of utterances. Across all participants, the total number of target words was 1,536. We excluded 126 targets (8%) due to caregivers' disfluencies or use of incorrect labels. As in Experiment 1, word durations for Experiment 2 were log-transformed, mean-centered, and scaled to correct for positive skew. Nontransformed results for Experiment 2 are depicted in Figure 3.

As in Experiment 1, we constructed maximal mixed-effects models with random slopes and intercepts by participants and items. We included audience (ADS or CDS), utterance position (first or second), and target word status (same referent or different referent) as factors. If a particular maximal model did not converge, we included the model that best fit the data. The best-fitting model included an Utterance Position by Target Word Status interaction and an Audience random slope by Participant. It also included Audience by Utterance Position interaction and a Target Word Status random slope by Target Label. Output for all models can be found in Appendix B.

We again began by constructing a model for all duration data across both ADS and CDS. Using this model, we found evidence

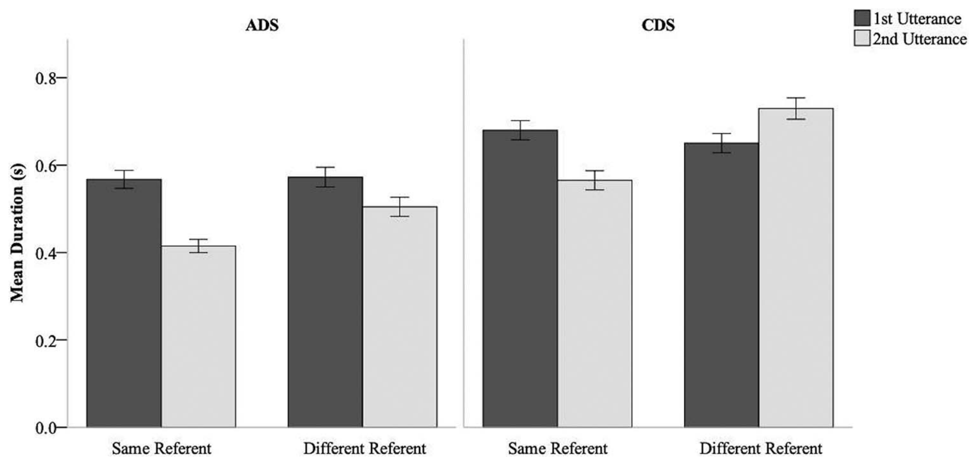

Figure 3. Target word durations, in seconds, for first and second utterances by audience and target word status. ADS = adult-directed speech; CDS = child-directed speech.

that targets in second utterance positions ($M = 0.55$ s, $SE = 0.03$ s) were shorter overall than targets in first utterance positions ($M = 0.62$ s, $SE = 0.02$ s), $t = -6.40$, $p < .0001$. We also find evidence that CDS ($M = 0.66$ s, $SE = 0.03$ s) was reliably longer than ADS ($M = 0.51$ s, $SE = 0.02$ s), $t = 5.97$, $p < .0001$. Finally, we found a significant three-way interaction between utterance position, audience, and target word status, $t = 3.71$, $p = .0002$. We constructed additional models to resolve our significant three-way interaction.[2]

We generated separate models for ADS and CDS target durations. For ADS, we found that targets in second utterance positions ($M = 0.60$ s, $SE = 0.02$ s) were shorter than targets in first utterance positions ($M = 0.50$ s, $SE = 0.02$ s), $t = -6.67$, $p < .0001$. However, an Utterance Position by Target Word Status interaction revealed more reduction of repeated information, $t = 3.97$, $p = .001$. Planned comparisons revealed that second mentions on same referent trials ($M = 0.41$ s, $SE = 0.02$ s) were shorter than first mentions ($M = 0.58$ s, $SE = 0.02$ s), $t = 6.73$, $p < .0001$. In addition, targets in second utterance positions ($M = 0.50$ s, $SE = 0.02$ s) were shorter than targets in first utterance positions ($M = 0.57$ s, $SE = 0.02$ s) for new trials, $t = 2.53$, $p = .03$. Speakers reduced when producing pairs of new referents as well as pairs of repeated referents. However, they reduced more for repeated information.

Our model for CDS revealed that second mentions of a target ($M = 0.64$ s, $SE = 0.03$ s) were reliably shorter than first mentions of a target ($M = 0.66$ s, $SE = 0.02$ s), $t = -4.45$, $p = .003$. We also found that speakers reduced second utterance targets in same referent pairs but not in different referent pairs, $t = 7.78$, $p < .0001$. Planned comparisons revealed reduction for same referent pairs: targets in second utterance positions ($M = 0.56$ s, $SE = 0.02$ s) were shorter than targets in first utterance positions ($M = 0.68$ s, $SE = 0.02$ s), $t = 5.11$, $p = .0003$. For different referent pairs, on the other hand, second mentions ($M = 0.73$ s, $SE = 0.02$ s) were longer than first mentions ($M = 0.65$ s, $SE = 0.02$ s), $t = -2.63$, $p = .03$.

Finally, we constructed a model comparing word durations for same referent trials across ADS and CDS. This model revealed an interaction between audience and utterance position, $t = 3.17$, $p = .005$. This was due to a greater amount of reduction of second utterance targets in ADS same referent pairs than in CDS same referent pairs, as seen in planned comparisons mentioned earlier. This replicates the results of Experiment 1. In other words, participants reduced repeated information more when speaking to adults than when speaking to children even if producing low-frequency labels for uncommon objects.

Discussion

In line with Experiment 1 and previous work (e.g., Bard et al., 2000; Bell et al., 2009; Fisher & Tokura, 1995; Fowler & Housum, 1987; see Wagner & Watson, 2010), speakers reduced repeated mentions of words. Importantly, they did so in both ADS and CDS, which replicates findings in both Experiment 1 and Fisher and Tokura (1995). However, the amount of reduction was greater for ADS than CDS, which replicates Experiment 1 but not Fisher and Tokura (1995). Finally, we found that speakers in Experiment 2 reliably lengthened targets overall in CDS compared to ADS, as in Fisher and Tokura (1995). The emergence of this effect in Experiment 2 is likely due to the stimuli: All target nouns in Experiment 2 were lower frequency than those in Experiment 1. Thus, children were unlikely to know them, which may have encouraged speakers to engage in CDS-related target lengthening. Because the words were unfamiliar, it is unclear if parents viewed the task as an opportunity to teach children an unknown word or to expose children to a less frequent word. For the purposes of our theoretical question, this distinction is not important, although future work may address this possibility more directly.

Experiment 2 provides compelling support for the idea that repetition reduction is a feature of speakers' production systems (e.g., Arnold & Watson, 2015; Bell et al., 2009; Kahn & Arnold,

[2] As in Experiment 1, we constructed a model that also included Global Mention (first, second, third, or fourth appearance of a target within the instructions). Here, there was a significant main effect of global mention, $t = 3.60$, $p < .001$. Target lengths tended to decrease between first ($M = 0.63$ s), second ($M = 0.58$ s), third ($M = 0.54$ s), and fourth ($M = 0.56$ s) mention. Importantly, however, all other critical effects were still present.

2012, 2015) rather than a product of audience design (e.g., Aylett & Turk, 2004; Jaeger, 2010; Lieberman, 1963). As in Experiment 1, speakers reduced repeated information regardless of their audience. This is even more striking here, because children were likely unfamiliar with the low-frequency nouns. Despite the additional needs of the audience, speakers still reduced repeated references. This provides general support for the idea that speakers may not actively control whether they engage in word-specific, finer-grained production processes (e.g., Fisher & Tokura, 1995).

However, we found evidence that runs counter to Fisher and Tokura's (1995) claim. In both Experiments 1 and 2, speakers reduced less when addressing a child than when addressing an adult, which suggests some level of audience design. Even though speakers may not have been able to actively control whether or not they engage in certain production processes, they did seem able to accommodate their audience, at least to some extent. Essentially, these results provide a nuanced account of how automated production processes and explicit audience design interact. Further work will be needed to understand the locus of the difference between our data and those of Fisher and Tokura (1995), which could have emerged from differences in utterance composition rather than features of ADS and CDS (e.g., Ratner, 1986; Fisher & Tokura, 1995; see Wagner & Watson, 2010).

In addition to the patterns of reduction in CDS and ADS mentioned above, we also found significant lengthening of newly referenced targets in CDS. This lengthening could be a feature of contrastive stress, which is used by English speakers to contrast a newly mentioned word from something previously said (e.g., Bolinger, 1961; Breen, Fedorenko, Wagner, & Gibson, 2010). In addition, there is some evidence that newly referenced words tend to receive emphasis in naturalistic CDS (e.g., Fernald & Mazzie, 1991), but past studies did not specifically examine words across first and second mentions. Importantly, we find the opposite pattern in ADS: Speakers reduced second targets in the new condition in ADS. This could be due to differences in utterance composition. Targets occurred at the end of first utterances and in the middle of second utterances. Utterance-final words tend to be lengthened (e.g., Ratner, 1986; Fisher & Tokura, 1995), which may explain our pattern of ADS results. Regardless, it is possible that this potential contrastive stress in Experiment 2 is another feature of audience design that speakers use when engaging in CDS.

General Results

We were also interested in comparing durations across Experiments 1 and 2 to test how common versus uncommon target nouns affected reduction. We rescaled and mean-centered raw word durations from both experiments to account for potential intrinsic differences in word length between targets in the two experiments. We constructed maximal models with random slopes and intercepts by participants and items. We included audience (ADS or CDS), utterance position (first or second), target word status (same referent or different referent), and experiment (1 or 2) as factors. Because the maximal model would not converge, we constructed models with different combinations of random slopes and used the best-fitting model for all analyses. The best-fitting model included an Audience by Utterance Position interaction and a Target Word Status random slope by Participant. It also included a Target Word Status by Utterance Position interaction and an Audience random slope by the Target Label. All coefficients, t-statistics, and p values can be found in Appendix B.

We found a main effect of experiment; target durations in Experiment 1 ($M = 0.43$ s, $SE = 0.03$ s) were shorter overall than those in Experiment 2 ($M = 0.58$ s, $SE = 0.04$ s), $t = 5.59$, $p < .0001$. We also found a significant interaction between experiment and audience, $t = 2.94$, $p = .004$. This interaction was due to the main effect of audience we found in Experiment 2, where CDS was reliably longer than ADS, and to the lack of a main effect of audience in Experiment 1. Specifically, the average difference in word duration between ADS and CDS was smaller in Experiment 1 than in Experiment 2, reflecting an overall lengthening of targets in CDS when targets were low frequency. Finally, we found a significant three-way interaction between audience, utterance position, and target word status, $t = 2.69$, $p = .007$. This interaction was present in both experiments and reveals that speakers reduced less in CDS than ADS.

General Discussion

Our results provide partial support for the claim that speakers have broad control over speech rate but do not have active control over finer-grained processes (e.g., Fisher & Tokura, 1995; see also Horton & Keysar, 1996 and Hwang, Brennan, & Huffman, 2015 for similar arguments). The fact that speakers reduced repeated mentions of targets when speaking to both adults and children suggests that there is some level of automaticity to this production process. Despite CDS being a clear instance of audience design that benefits young children's understanding (e.g., Liu et al., 2003; Schwab & Lew-Williams, 2016b; Thiessen et al., 2005), speakers reduced repeated referents when talking to children. Importantly, they also do so under conditions that should elicit as much audience design as possible, namely when their listeners are following instructions (e.g., Yoon et al., 2012). It seems, then, that some feature inherent in speakers' production systems is driving this tendency (e.g., Arnold & Watson, 2015; Bell et al., 2009; Kahn & Arnold, 2012, 2015).

Critically, however, speakers clearly engaged in elements of audience design. Most obviously in Experiment 2, CDS was longer than ADS, which demonstrates speakers' recognition that children may have needed additional support in order to understand low-frequency nouns. More importantly, if reduction was a truly automatic process, we should have found the same amount of reduction across both ADS and CDS. We did not. Instead, we found that speakers reduced less in CDS than ADS across both experiments. This runs counter to Fisher and Tokura's (1995) finding that there was more reduction in CDS than ADS. Their finding, though, could have been due to their unscripted task, which led to differences between CDS and ADS in terms of utterance composition (e.g., Ratner, 1986; Fisher & Tokura, 1995; see Wagner & Watson, 2010). Using more controlled sentences across both speech registers, we find that speakers are able to control the amount that they reduce, despite engaging in automatic repetition reduction. Importantly, they reduce less for children than adults, suggesting a complex interplay between audience design and automated production processes. Our results imply that speakers may engage in as much audience design as possible, but this is not sufficient to overcome constraints of the production system that ultimately lead to repetition reduction.

Other research exploring CDS offers support for the claims we make here. In a study on word stress and repetition by Bortfeld and Morgan (2010), adults seemed to "reset" word lengths every other utterance, suggesting that they cyclically emphasize words within the context of a conversation with their infants. This provides additional evidence that some automatic production constraints are likely at play, but speakers maintain an ability to modulate certain features that may aid listener comprehension. Specifically, speakers may reduce a referring expression if it is primed by recent mentions (i.e., a production constraint), but concurrently reset the pitch accent of the same referring expression to recapture infant attention. These ideas call for further scrutiny of the dynamics between various features of audience design.

Another perspective on these data is that reduction of repeated words can sometimes serve as a helpful cue for listener comprehension, as it signals words that have already been used in the discourse (e.g., Isaacs & Watson, 2010). Indeed, young children are able to learn reduced forms of words that occur commonly in CDS (e.g., Lahey & Ernestus, 2014). In addition, 7.5-month-old infants are better at comprehending unstressed words if they are first exposed to those words in unstressed forms (Bortfeld & Morgan, 2010). If repetition reduction is an informative cue, one might expect speakers to reduce more in CDS in order to scaffold comprehension of discourse-given words (as was the case in Fisher & Tokura, 1995). However, the fact that speakers in our study reduced CDS less than ADS complicates this idea. This suggests, then, that a potentially more complicated version of audience design is at play. It seems that speakers could be navigating the utility of reduction while also recognizing that children need some scaffolding. Of course, this could differ across age of a listener. In our study, parents were speaking to their preschoolers, and in previous research parents were speaking to 14-month-old infants (Fisher & Tokura, 1995). However, there are no existing data suggesting that such differences would fundamentally alter the patterns of results across the experiments. Additional research is needed to ascertain whether there are differences in caregivers' word-level audience design when communicating with infants versus young children.

The fact that we observed less reduction in CDS than ADS warrants interpretation of why speakers treat the two groups of listeners differently. In particular, it is possible that speakers reduce repeated words less in CDS because they are attempting to maximize the intelligibility of the word while still providing acoustic cues to the discourse status of the word. Consequently, words are reduced, but not as much as they would be in ADS, where the adult audience needs less help in decoding what is new and old in the speech signal. Relatedly, speakers may have been less invested in the task with adults, because they recognized the adult listener's competence. This lack of investment could have resulted in differences in reduction we observed. Indeed, other work has demonstrated that speaker engagement can influence some prosodic features (e.g., Buxó-Lugo, Toscano, & Watson, 2018). Regardless, future work will need to directly address whether repetition reduction is a reliable tool for comprehension—especially in children—and whether the extent to which parents reduce varies across early development. In particular, it is interesting to consider whether adult speakers are engaging in certain linguistic behaviors—lesser reduction in CDS, for example—that may not maximize young children's entry into language learning and comprehension.

In addition to the reduction mentioned above, we also found that speakers lengthened second mentions of newly referenced words in CDS in Experiment 2. This could be a feature of contrastive emphasis used to scaffold children's understanding of instructions involving uncommon objects. Somewhat relatedly, our cross-experiment analysis revealed that speakers produced unfamiliar labels with longer duration than familiar labels. This could be due to the fact that content words are shorter if they are more frequent (Bell et al., 2009; Zipf, 1936), and the words used in Experiment 1 were more frequent than those in Experiment 2. Nevertheless, the fact that we found that there was an interaction between audience and experiment suggests that speakers engaged in some level of audience design: While there was a general increase in duration for unfamiliar labels compared to familiar ones, this lengthening was greater in CDS than in ADS. Whether this is due to word frequency or complexity is not relevant. What is relevant is the fact that caregivers engaged in some level of audience design when reading more difficult instructions to children.

Another potentially problematic aspect of the design was that we allowed caregivers to point to unfamiliar objects in Experiment 2. This served to help enable natural interaction, but caregivers were not explicitly given permission to do so in Experiment 1. An ostensive gesture like pointing may influence a speaker's tendency to engage in repetition reduction, because there is an additional cue to the speaker's intended referent. That is, pointing could stand in for elongation or reduction by helping to disambiguate a speaker's intended referent. However, our results do not support this possibility. We found similar attenuation of reduction in CDS in both Experiments 1 and 2, despite the fact that caregivers were only told to point in Experiment 2. Moreover, speakers lengthened new mentions of targets in Experiment 2, which would not be expected if pointing served to reduce audience design. Thus, it is unlikely that pointing drove the effects we observed across our two experiments, but the potential interaction between gesture and audience design is a promising direction for future research.

One feature of our experimental manipulation is that parents completed ADS trials before CDS trials. This feature of the design is unlikely to undermine the validity of the results for two reasons. First, if audience order influenced word duration, the current design should have reduced the likelihood of finding the observed effects. That is, if participants were systematically reducing durations simply because they were reading the same targets repeatedly, we should have found evidence that CDS trials were reliably shorter than ADS trials, either overall or for second mentions. We did not find this. Instead, we found lengthening of CDS in Experiment 2, plus no statistical difference between ADS and CDS in Experiment 1. Second, the current design served a needed pragmatic function. Parents were told that they would practice the task with an experimenter first, which enabled us to keep participants blind to the actual manipulation of interest (ADS vs. CDS). That is, allowing participants to conduct the task with an adult first provided a natural cover story. The reverse order would be pragmatically unworkable and would introduce unwanted artificiality to their speech with the adult experimenter.

However, it is worth noting that using a confederate as an addressee may be problematic for detecting reduction effects (see Kuhlen & Brennan, 2013 for additional discussion). Indeed, in our experiment the confederate did not need comprehension scaffolding because she knew the design of the experiment, something adult speakers likely recognized. However, we saw repetition reduction even during communication with a knowledgeable adult. This offers evidence that reduction is, on some level, the result of a speaker's own internal production constraints. Critically, though, the overall rate of caregivers' word-level reduction did differ when speaking to a naïve child versus a knowledgeable adult. This leaves open the question of whether or how speakers engage in fine-grained modulation of speech based on the perceived knowledge of their listeners.

Our results are relevant for constraining information theoretic explanations of reduction (e.g., Aylett & Turk, 2004; Jaeger, 2010). One ongoing question in current research is understanding which contextual factors make words more versus less predictable, and, therefore, more versus less informative. In many implementations of information theoretic accounts, information density is calculated based on the predictability of a word or structure given some linguistic context, such as the probability of the word in a syntactic parse or the probability of a word given the preceding word string. In the case of repeated words, words are shortened because they are linguistically more predictable, which maintains information density, and this facilitates processing for the listener. However, the current data suggest that contextual factors such as the likely knowledge state of the audience can influence word duration. For information theoretic models to account for these types of data, what counts as contextual information must include situational, nonlinguistic contexts (such as the range of listeners' familiarity with words or conversation topics), which are more difficult to formalize than linguistic contexts. More work is needed to understand how situational contexts influence predictability. Nevertheless, it seems that there is a complicated interplay between a number of competing factors, including word position, repetition, and communicative goals.

It is important to note that these results do not provide insight into exactly which feature of the production system drives reduction. The goal of the current study was to ask a more general question: is repetition reduction driven by an internal production mechanism or by an externally imposed consideration of audience? Our findings support the claim that reduction may in part stem from an internal mechanism (e.g., Arnold & Watson, 2015; Bell et al., 2009; Kahn & Arnold, 2012, 2015). Other lines of research have explored specific, speaker-centered cognitive mechanisms underlying speech production (e.g., Jacobs, Yiu, Watson, & Dell, 2015; Watson, Buxó-Lugo, & Simmons, 2015; Yiu & Watson, 2015). There is likely an interaction between audience design and automatic processes, such that speakers may simply represent a conversational partner with a minimal cue, such as whether or not they are a child (Brennan, Galati, & Kuhlen, 2010; Galati & Brennan, 2010), and this simple representation could affect automatic production processes. This possibility is captured by computational models of perspective taking (e.g., Dale et al., 2018; Duran & Dale, 2014). Exploratory analyses of a particular target's first through fourth appearance across the experiments offer equivocal evidence supporting these points. Further research is needed to understand the precise time course of the interactions taking place within the production system.

Conclusion

In summary, by studying ADS versus CDS, we found an interplay between audience design and automatic processes during speech production. In particular, it seems that these may be nested processes. It is not clear from the present research, however, whether audience design is nested within an automatic production system or vice versa. Reduction of repeated referents may be a byproduct of speakers' production needs (e.g., Fisher & Tokura, 1995), but there is still room for speakers to tailor utterances based on their audience's needs, knowledge, or proficiency. Specifically, we provide confirmation that repetition reduction occurs in CDS using a controlled paradigm intended to elicit audience design.

These results provide new insight into a phenomenon originally investigated by Fisher and Tokura (1995), which proposed that speakers may not have control over certain production processes. However, our results complicate the claim that speakers do not have active control over fine-grained, word-specific durational decisions. On the contrary, speakers reduced repeated mentions of targets less when speaking to children than when speaking to adults. Thus, there is likely to be a complex relation between ease of retrieval for a speaker and ease of comprehension for a listener, and listeners may be attuned to patterns of speech that are related to speakers' default production tendencies. This attunement makes it challenging to tease apart conflicting accounts of production, but the present experiments offer insights into debates about how production and comprehension interact across communicative contexts (e.g., Bard & Aylett, 2004; Galati & Brennan, 2010). In addition, our results contribute to a broader literature on the interplay between automatic and contextual factors in human cognition, which has been examined in, for example, expression of racial prejudice (Lowery, Hardin, & Sinclair, 2001) and memory (Hasher & Zacks, 1979). Our results suggest that, in human communication, automatic processes could be tempered by contextual features, such as the comprehension needs of a young child.

References

Arnold, J. E., Kahn, J. M., & Pancani, G. C. (2012). Audience design affects acoustic reduction via production facilitation. *Psychonomic Bulletin & Review, 19,* 505–512. http://dx.doi.org/10.3758/s13423-012-0233-y

Arnold, J. E., & Watson, D. G. (2015). Synthesizing meaning and processing approaches to prosody: Performance matters. *Language, Cognition and Neuroscience, 30,* 88–102. http://dx.doi.org/10.1080/01690965.2013.840733

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. http://dx.doi.org/10.1177/00238309040470010201

Bard, E. G., & Anderson, A. H. (1983). The unintelligibility of speech to children. *Journal of Child Language, 10,* 265–292. http://dx.doi.org/10.1017/S0305000900007777

Bard, E. G., 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. http://dx.doi.org/10.1006/jmla.1999.2667

Bard, E. G., & Aylett, M. P. (2004). Referential form, word duration, and modeling the listener in spoken dialogue. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 173–191). Cambridge, MA: MIT Press.

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. http://dx.doi.org/10.1016/j.jml.2008.06.003

Boersma, P., & Weenink, D. (2017). Praat: Doing phonetics by computer (Version 6.0.37) [Computer software]. Retrieved from http://www.praat.org/

Bolinger, D. L. (1961). Contrastive accent and contrastive stress. *Language, 37,* 83–96. http://dx.doi.org/10.2307/411252

Bortfeld, H., & Morgan, J. L. (2010). Is early word-form processing stress-full? How natural variability supports recognition. *Cognitive Psychology, 60,* 241–266. http://dx.doi.org/10.1016/j.cogpsych.2010.01.002

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes, 25,* 1044–1098. http://dx.doi.org/10.1080/01690965.2010.504378

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation, 53,* 301–344. http://dx.doi.org/10.1016/S0079-7421(10)53008-1

Brown-Schmidt, S., & Heller, D. (2018). Perspective taking during conversation. In G. Gaskell & S. A. Rueschemeyer (Eds.), *Oxford handbook of psycholinguistics*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780198786825.013.23

Buxó-Lugo, A., Toscano, J. C., & Watson, D. G. (2018). Effects of participant engagement on prosodic prominence. *Discourse Processes, 55,* 305–323. http://dx.doi.org/10.1080/0163853X.2016.1240742

Chafe, W. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse* (pp. 21–51). Amsterdam, the Netherlands: John Benjamins. http://dx.doi.org/10.1075/tsl.11.03cha

Chambers, C. G., & Smyth, R. (1998). Structural parallelism and discourse coherence: A test of centering theory. *Journal of Memory and Language, 39,* 593–608. http://dx.doi.org/10.1006/jmla.1998.2575

Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61,* 1584–1595. http://dx.doi.org/10.2307/1130766

Cristia, A. (2013). Input to language: The phonetics and perception of infant-directed speech. *Language and Linguistics Compass, 7,* 157–170. http://dx.doi.org/10.1111/lnc3.12015

Dale, R., Galati, A., Alviar, C., Contreras Kallens, P., Ramirez-Aristizabal, A. G., Tabatabaeian, M., & Vinson, D. W. (2018). Interacting timescales in perspective-taking. *Frontiers in Psychology, 9,* 1278. http://dx.doi.org/10.3389/fpsyg.2018.01278

Duran, N. D., & Dale, R. (2014). Perspective-taking in dialogue as self-organization under social constraints. *New Ideas in Psychology, 32,* 131–146. http://dx.doi.org/10.1016/j.newideapsych.2013.03.004

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development, 8,* 181–195. http://dx.doi.org/10.1016/S0163-6383(85)80005-9

Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology, 27,* 209–221. http://dx.doi.org/10.1037/0012-1649.27.2.209

Fisher, C., & Tokura, H. (1995). The given-new contract in speech to infants. *Journal of Memory and Language, 34,* 287–310. http://dx.doi.org/10.1006/jmla.1995.1013

Fowler, C. A., & 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, 26,* 489–504. http://dx.doi.org/10.1016/0749-596X(87)90136-7

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language, 62,* 35–51. http://dx.doi.org/10.1016/j.jml.2009.09.002

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388. http://dx.doi.org/10.1037/0096-3445.108.3.356

Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition, 59,* 91–117. http://dx.doi.org/10.1016/0010-0277(96)81418-1

Hwang, J., Brennan, S. E., & Huffman, M. K. (2015). Phonetic adaptation in non-native spoken dialogue: Effects of priming and audience design. *Journal of Memory and Language, 81,* 72–90. http://dx.doi.org/10.1016/j.jml.2015.01.001

Isaacs, A. M., & Watson, D. G. (2010). Accent detection is a slippery slope: Direction and rate of F0 change drives listeners' comprehension. *Language and Cognitive Processes, 25,* 1178–1200. http://dx.doi.org/10.1080/01690961003783699

Jacobs, C. L., Yiu, L. K., Watson, D. G., & Dell, G. S. (2015). Why are repeated words produced with reduced durations? Evidence from inner speech and homophone production. *Journal of Memory and Language, 84,* 37–48. http://dx.doi.org/10.1016/j.jml.2015.05.004

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology, 61,* 23–62. http://dx.doi.org/10.1016/j.cogpsych.2010.02.002

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. *Typological Studies in Language, 45,* 229–254.

Kahn, J. M., & Arnold, J. E. (2012). A processing-centered look at the contribution of givenness to durational reduction. *Journal of Memory and Language, 67,* 311–325. http://dx.doi.org/10.1016/j.jml.2012.07.002

Kahn, J. M., & Arnold, J. E. (2015). Articulatory and lexical repetition effects on durational reduction: Speaker experience vs. common ground. *Language, Cognition and Neuroscience, 30,* 103–119. http://dx.doi.org/10.1080/01690965.2013.848989

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: When confederates might be hazardous to your data. *Psychonomic Bulletin & Review, 20,* 54–72. http://dx.doi.org/10.3758/s13423-012-0341-8

Lahey, M., & Ernestus, M. (2014). Pronunciation variation in infant-directed speech: Phonetic reduction of two highly frequent words. *Language Learning and Development, 10,* 308–327. http://dx.doi.org/10.1080/15475441.2013.860813

Lam, T. Q., & Watson, D. G. (2010). Repetition is easy: Why repeated referents have reduced prominence. *Memory & Cognition, 38,* 1137–1146. http://dx.doi.org/10.3758/MC.38.8.1137

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech, 6,* 172–187. http://dx.doi.org/10.1177/002383096300600306

Liu, H. M., Kuhl, P. K., & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science, 6*(3), F1–F10. http://dx.doi.org/10.1111/1467-7687.00275

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology, 81,* 842–855. http://dx.doi.org/10.1037/0022-3514.81.5.842

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk, Vol. 2: The database* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Pate, J. K., & Goldwater, S. (2015). Talkers account for listener and channel characteristics to communicate efficiently. *Journal of Memory and Language, 78,* 1–17. http://dx.doi.org/10.1016/j.jml.2014.10.003

Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior & Development, 15,* 325–345. http://dx.doi.org/10.1016/0163-6383(92)80003-D

Piazza, E. A., Iordan, M. C., & Lew-Williams, C. (2017). Mothers consistently alter their unique vocal fingerprints when communicating with infants. *Current Biology, 27,* 3162–3167.e3. http://dx.doi.org/10.1016/j.cub.2017.08.074

Prince, E. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical pragmatics* (pp. 223–255). New York, NY: Academic Press.

Ratner, N. B. (1986). Duration cues which mark clause boundaries in mother-child speech. *Journal of Phonetics, 14,* 303–309. http://dx.doi.org/10.1016/S0095-4470(19)30670-9

Scarborough, R., Dmitrieva, O., Hall-Lew, L., Zhao, Y., & Brenier, J. (2007). An acoustic study of real and imagined foreigner-directed speech. *The Journal of the Acoustical Society of America, 121,* 3044. http://dx.doi.org/10.1121/1.4781735

Schwab, J. F., & Lew-Williams, C. (2016a). Language learning, socioeconomic status, and child-directed speech. *WIREs Cognitive Science, 7,* 264–275. http://dx.doi.org/10.1002/wcs.1393

Schwab, J. F., & Lew-Williams, C. (2016b). Repetition across successive sentences facilitates young children's word learning. *Developmental Psychology, 52,* 879–886. http://dx.doi.org/10.1037/dev0000125

Schwab, J. F., Rowe, M. L., Cabrera, N., & Lew-Williams, C. (2018). Fathers' repetition of words is coupled with children's vocabularies. *Journal of Experimental Child Psychology, 166,* 437–450. http://dx.doi.org/10.1016/j.jecp.2017.09.012

Smith, C. (2007). Prosodic accommodation by French speakers to a non-native interlocutor. *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1081–1084). Retrieved from https://pdfs.semanticscholar.org/0012/f7576aa59432ecfd2fc588707d60872e0a3a.pdf

Smyth, R. (1994). Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research, 23,* 197–229. http://dx.doi.org/10.1007/BF02139085

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review, 27,* 501–532. http://dx.doi.org/10.1016/j.dr.2007.06.002

Stalnaker, R. (1973). Presuppositions. *Journal of Philosophical Logic, 2,* 447–457. http://dx.doi.org/10.1007/BF00262951

Thiessen, E. D., Hill, E. A., & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy, 7,* 53–71. http://dx.doi.org/10.1207/s15327078in0701_5

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech. *Speech Communication, 49,* 2–7. http://dx.doi.org/10.1016/j.specom.2006.10.003

Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes, 25,* 905–945. http://dx.doi.org/10.1080/01690961003589492

Watson, D. G., Buxó-Lugo, A., & Simmons, D. C. (2015). The effect of phonological encoding on word duration: Selection takes time. In L. Frazier & E. Gibson (Eds.), *Explicit and implicit prosody in sentence processing* (pp. 85–98). Cham, Switzerland: Springer. http://dx.doi.org/10.1007/978-3-319-12961-7_5

Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue Canadienne de Psychologie, 43,* 230–246. http://dx.doi.org/10.1037/h0084224

Werker, J. F., Pegg, J. E., & McLeod, P. J. (1994). A cross-language investigation of infant preference for infant-directed communication. *Infant Behavior & Development, 17,* 323–333. http://dx.doi.org/10.1016/0163-6383(94)90012-4

Xu, N., Burnham, D., Kitamura, C., & Vollmer-Conna, U. (2013). Vowel hyperarticulation in parrot-, dog-and infant-directed speech. *Anthrozoös, 26,* 373–380.

Yiu, L. K., & Watson, D. G. (2015). When overlap leads to competition: Effects of phonological encoding on word duration. *Psychonomic Bulletin & Review, 22,* 1701–1708. http://dx.doi.org/10.3758/s13423-015-0828-1

Yoon, S. O., Koh, S., & Brown-Schmidt, S. (2012). Influence of perspective and goals on reference production in conversation. *Psychonomic Bulletin & Review, 19,* 699–707. http://dx.doi.org/10.3758/s13423-012-0262-6

Zipf, G. K. (1936). *The psycho-biology of language: An introduction to dynamic philology*. New York, NY: Routledge.

Appendix A

List of Instructions for Experiments 1 and 2

Instructions Used in Experiment 1

Get the dog. Can you put it in the little wooden house?
Now put the dog on top of the pink castle.

Get the cat. Can you put it in the big ball?
Now put the duck in the purple castle.

Get the cow. Can you put it under the big log?
Now put the cow on top of the big wooden house.

Get the deer. Can you put it in the big wooden house?
Now put the pig on top of the big log.

Get the duck. Can you put it in the little ball?
Now put the duck in the pink castle.

Get the goat. Can you put it on top of the little wooden house?
Now put the dog on top of the big log.

Get the cat. Can you put it in the pink castle?
Now put the cat on top of the little log.

Get the dog. Can you put it in the little wooden house?
Now put the bird in the little ball.

Get the deer. Can you put it in the big ball?
Now put the deer in the purple castle.

Get the cow. Can you put it under the big log?
Now put the goat on top of the purple castle.

Get the pig. Can you put it in the big ball?
Now put the pig on top of the little log.

Get the duck. Can you put it in the little ball?
Now put the cat on top of the purple castle.

Get the bird. Can you put it under the little log?
Now put the bird in the big ball.

Get the pig. Can you put it in the little ball?
Now put the cow on top of the big wooden house.

Get the goat. Can you put it on top of the little wooden house?
Now put the goat on top of the pink castle.

Get the bird. Can you put it under the little log?
Now put the deer in the big wooden house.

Instructions Used in Experiment 2

Get the thimble. Can you put it in the little wooden house?
Now put the thimble on top of the pink castle.

Get the spade. Can you put it in the big ball?
Now put the globe in the purple castle.

Get the felt. Can you put it under the big log?
Now put the felt on top of the big wooden house.

Get the spool. Can you put it in the big wooden house?
Now put the twine on top of the big log.

Get the globe. Can you put it in the little ball?
Now put the globe in the pink castle.

Get the brooch. Can you put it on top of the little wooden house?
Now put the thimble on top of the big log.

Get the spade. Can you put it in the pink castle?
Now put the spade on top of the little log.

Get the thimble. Can you put it in the little wooden house?
Now put the gnome in the little ball.

Get the spool. Can you put it in the big ball?
Now put the spool in the purple castle.

Get the felt. Can you put it under the big log?
Now put the brooch on top of the purple castle.

Get the twine. Can you put it in the big ball?
Now put the twine on top of the little log.

Get the globe. Can you put it in the little ball?
Now put the spade on top of the purple castle.

Get the gnome. Can you put it under the little log?
Now put the gnome in the big ball.

Get the twine. Can you put it in the little ball?
Now put the felt on top of the big wooden house.

Get the brooch. Can you put it on top of the little wooden house?
Now put the brooch on top of the pink castle.

Get the gnome. Can you put it under the little log?
Now put the spool in the big wooden house.

(*Appendices continue*)

Appendix B

Model Output for All Reported Analyses

Table B1

Output for Experiment 1 Total Model

Predictor	Estimate (beta)	*t* value	*p* value
Intercept	0.25	1.95	.06
Audience	0.12	1.08	.29
Utterance position	−1.17	−7.94	<.0001*
Target word status	−0.02	−0.25	.81
Target Word Status × Utterance Position	0.53	4.16	.0002*
Target Word Status × Audience	0.06	0.49	.62
Audience × Utterance Position	0.38	3.78	.0001*
Audience × Utterance Position × Target Word Status	0.37	2.38	.02*

Note. R syntax: Lmer[Scale(log[Duration]) ~ Target Word Status × Audience × Utterance Position + (1 + Condition × Utterance Position | Participant) + (1 + Audience | Participant) + (1 + Target Word Status × Utterance Position | Label) + (1 + Audience | Label)].
* $p < .05$.

Table B2

Output for Experiment 1 ADS Model

Predictor	Estimate (beta)	*t* value	*p* value
Intercept	0.49	3.33	.003*
Target word status	−0.02	−0.31	.76
Utterance position	−1.24	−6.78	<.0001*
Target Word Status × Utterance Position	0.56	3.81	.002*

Note. ADS = adult-directed speech. R syntax: Lmer[Scale(log[Duration]) ~ Target Word Status × Utterance Position + (1 + Target Word Status × Utterance Position | Participant) + (1 + Target Word Status × Utterance Position | Label)].
* $p < .05$.

Table B3

Output for Experiment 1 CDS Model

Predictor	Estimate (beta)	*t* value	*p* value
Intercept	0.12	0.86	.40
Target word status	0.05	0.35	.73
Utterance position	−0.80	−5.26	<.0001*
Target Word Status × Utterance Position	0.90	5.51	<.0001*

Note. CDS = child-directed speech. R syntax: Lmer[Scale(log[Duration]) ~ Target Word Status × Utterance Position + (1 + Target Word Status × Utterance Position | Participant) + (1 + Target Word Status × Utterance Position | Label)].
* $p < .05$.

(*Appendices continue*)

Table B4
Output for Experiment 1 Same Referent Trials Model

Predictor	Estimate (beta)	t value	p value
Intercept	0.37	2.86	.01*
Audience	0.12	1.01	.33
Utterance position	−1.14	−7.07	<.0001*
Audience × Utterance Position	0.38	2.39	.04*

Note. R syntax: Lmer[Scale(log[Duration]) ~ Utterance Position × Audience + (1 + Utterance Position × Audience | Participant) + (1 + Utterance × Audience | Label)].
* $p < .05$.

Table B5
Output for Experiment 2 Total Model

Predictor	Estimate (beta)	t value	p value
Intercept	−0.01	−0.07	.95
Audience	0.54	5.97	<.0001*
Utterance position	−1.02	−6.40	<.0001*
Target word status	0.02	0.29	.78
Target Word Status × Utterance Position	0.57	4.86	<.0001*
Target Word Status × Audience	−0.16	−1.82	.07
Audience × Utterance Position	0.37	3.81	.0003*
Audience × Utterance Position × Target Word Status	0.47	3.71	.0002*

Note. R syntax: Lmer[Scale(log[Duration]) ~ Target Word Status × Audience × Utterance Position + (1 + Audience | Participant) + (1 + Utterance Position × Target Word Status | Participant) + (1 + Target word status | Label) + (1 + Audience × Utterance Position | Label)].
* $p < .05$.

Table B6
Output for Experiment 2 ADS Model

Predictor	Estimate (beta)	t value	p value
Intercept	0.38	2.59	.02*
Target word status	0.02	0.24	.82
Utterance position	−1.09	−6.67	<.0001*
Target Word Status × Utterance Position	0.61	4.56	.0007*

Note. ADS = adult-directed speech. R syntax: Lmer[Scale(log[Duration]) ~ Target Word Status × Utterance Position + (1 + Target Word Status × Utterance Position | Participant) + (1 + Target Word Status × Utterance Position | Label)].
* $p < .05$.

(*Appendices continue*)

Table B7
Output for Experiment 2 CDS Model

Predictor	Estimate (beta)	*t* value	*p* value
Intercept	0.15	0.86	.40
Target word status	−0.16	−1.95	.06
Utterance position	−0.73	−4.45	.0003*
Target Word Status × Utterance Position	1.16	7.78	<.0001*

Note. CDS = child-directed speech. R syntax: Lmer[Scale(log[Duration]) ~ Target Word Status × Utterance Position + (1 + Target Word Status × Utterance Position | Participant) + (1 + Target Word Status × Utterance Position | Label))].
* $p < .05$.

Table B8
Output for Experiment 2 Same Referent Trials Model

Predictor	Estimate (beta)	*t* value	*p* value
Intercept	0.16	1.16	.26
Audience	0.53	5.79	<.0001*
Utterance position	−0.99	−6.70	<.0001*
Audience × Utterance Position	0.35	3.17	.005*

Note. R syntax: Lmer[Scale(log[Duration]) ~ Utterance Position × Audience + (1 + Utterance Position × Audience | Participant) + (1 + Utterance Position × Audience | Label)].
* $p < .05$.

Table B9
Output for Cross-Experiment Model

Predictor	Estimate (beta)	*t* value	*p* value
Intercept	−0.25	−2.45	.02*
Experiment	0.65	5.59	<.0001*
Audience	0.15	1.92	.06
Utterance position	−1.13	−8.79	<.0001*
Target word status	−0.03	−0.39	.70
Target Word Status × Utterance Position	0.49	4.74	<.0001*
Target Word Status × Audience	0.03	0.25	.81
Target Word Status × Experiment	0.05	0.41	.68
Audience × Utterance Position	0.41	3.62	.0004*
Audience × Experiment	0.30	2.94	.004*
Utterance Position × Experiment	0.37	2.67	.01*
Audience × Utterance Position × Target Word Status	0.39	2.69	.007*
Audience × Utterance Position × Experiment	−0.20	−1.48	.14
Audience × Target Word Status × Experiment	−0.16	−1.13	.26
Target Word Status × Utterance Position × Experiment	0.03	0.18	.85
Target Word Status × Utterance Position × Audience × Experiment	0.02	0.08	.94

Note. R syntax: Lmer[Scale(log[Duration]) ~ Experiment × Audience × Target Word Status × Utterance Position + (1 + Target word status | Participant) + (1 + Audience × Utterance Position | Participant) + (1 + Audience | Label) + (1 + Target Word Status × Utterance Position | Label)].
* $p < .05$.

Received September 10, 2019
Revision received May 29, 2020
Accepted June 2, 2020 ■