

# Audience design affects acoustic reduction via production facilitation

Jennifer E. Arnold · Jason M. Kahn · Giulia C. Pancani

Published online: 15 March 2012  
© Psychonomic Society, Inc. 2012

**Abstract** In this article, we examine the hypothesis that acoustic variation (e.g., reduced vs. prominent forms) results from audience design. Bard et al. (Journal of Memory and Language 42:1–22, 2000) have argued that acoustic prominence is unaffected by the speaker's estimate of addressee knowledge, using paradigms that contrast speaker and addressee knowledge. This question was tested in a novel paradigm, focusing on the effects of addressees' feedback about their understanding of the speaker's intended message. Speakers gave instructions to addressees about where to place objects (e.g., *the teapot goes on red*). The addressee either anticipated the object, by picking it up before the instruction, or waited for the instruction. For anticipating addressees, speakers began speaking more quickly and pronounced the word *the* with shorter duration, demonstrating effects of audience design. However, no effects appeared on the head noun (e.g., *teapot*), as measured by duration, amplitude, and perceived intelligibility. These results are consistent with a mechanism in which evidence about addressee understanding facilitates production processes, as opposed to triggering particular acoustic forms.

**Keywords** Acoustic prominence · Reference · Language production · Audience design · Prosody

Spoken words vary in their degrees of acoustic prominence, or intelligibility. Words that are given (e.g., repeated) or

predictable in context tend to be acoustically reduced, whereas new or unpredictable words are often acoustically prominent. Acoustic reduction can be signaled by shorter duration, reduced pitch or pitch variation, reduced intensity, and lower intelligibility (Wagner & Watson, 2010).

Although such variation abounds, the mechanisms behind it are not fully understood. An unresolved question is whether the variation reflects audience design. Extensive evidence has shown that speakers select linguistic forms on the basis of the knowledge, intentions, or goals of their addressee (e.g., Brennan & Clark, 1996). However, the extent of addressee-oriented production is heavily debated, and some production processes proceed without reference to the addressee's knowledge (e.g., Ferreira & Dell, 2000; Horton & Keysar, 1996). Here we examine the role of audience design on acoustic prominence and consider how audience design relates to utterance planning.

An audience-design explanation states that acoustic variation serves the needs of the addressee. Acoustically prominent pronunciations are used when the word is harder to understand in context, as new or unpredictable words are (Lindblom, 1990; Wright, 2004; see Smiljanic & Bradlow, 2009). This proposal is related to the view that speakers follow pragmatic rules based on a word's information status, for example selecting reduced forms for topical, given, and predictable words, but prominent forms for new information (Halliday, 1967).

By contrast, the speaker-internal view of acoustic variation is based on the observation that reduction tends to happen in situations in which production is facilitated. Previously mentioned material is easier to mention again, because representations in the production system (e.g., conceptual, lexical, and articulatory) have recently been activated. Likewise, when a word is probable or supported by the context, its lexical representation is likely to receive an

---

J. E. Arnold · J. M. Kahn · G. C. Pancani  
Department of Psychology, University of North Carolina,  
Chapel Hill, NC, USA

J. E. Arnold (✉)  
CB #3270 Davie Hall, Rm. 337B,  
Chapel Hill, NC 27599-3270, USA  
e-mail: jarnold@email.unc.edu

activation boost (e.g., Balota et al., 1989; Bell, Brenier, Gregory, Girand, & Jurafsky, 2009). Referential predictability also matters: When a referent is expected, the conceptual representation is likely preactivated, and utterance planning can begin earlier (Kahn & Arnold, 2012; Watson, Arnold, & Tanenhaus, 2008).

Distinguishing these views is difficult, because the same contexts favor attenuation for both speaker and addressee. Indeed, information-theoretic theories suggest that reduction strikes a balance between the comprehension needs of the listener and the efficiency of the production system (Aylett & Turk, 2004).

Nevertheless, evidence suggests that acoustic prominence is driven by at least some speaker-internal processes. Kahn and Arnold (2012) found that the degree of reduction corresponded to the number of levels of production facilitation. Words were shortened when preceded by linguistic (spoken word) or nonlinguistic (flashing picture) primes. Critically, reduction was greater for the linguistic than the nonlinguistic prime, suggesting that the speaker's experience with the word facilitated pronunciation over and above the activation of the concept. Moreover, when addressee knowledge of givenness was manipulated, only the speaker's knowledge affected durations of the target word (Kahn & Arnold, 2011).

Similarly, Bard et al. (2000; Bard & Aylett, 2004) argued against the audience-design view with data from their map task corpus, in which reduction occurred even for new addressees, and only on the basis of the speaker's experience. They proposed the dual-process hypothesis, which states that "listener modeling is not obligatory. . . . If it takes place anywhere, it will affect the slower-cycling positional level rather than faster-cycling articulatory level" (Bard & Aylett, 2004, p. 178). This predicts that listener modeling will affect lexical choices (e.g., pronoun use), but not acoustic prominence.

However, these studies did not exclude the possibility that acoustic reduction may be partially addressee-driven. Galati and Brennan (2010) observed that Bard et al. (2000) only examined repeated mentions to new addressees, and Galati and Brennan hypothesized that repeated mentions to the same addressee should be even more reduced. When they tested this idea, they found no difference in word durations (Exp. 1), but listeners rated repeated mentions to the same addressee as less intelligible than those to a different addressee (Exp. 2).

In sum, evidence for audience-design effects is inconsistent across experiments, and there is no evidence to date that audience design affects word durations. One feature of previous studies is that they have featured paradigms in which speakers had to keep track of differences in knowledge between themselves and their addressee, or between different addressees. In the present study, we used a different approach to seek evidence of audience design on acoustic reduction: We examined audience feedback about

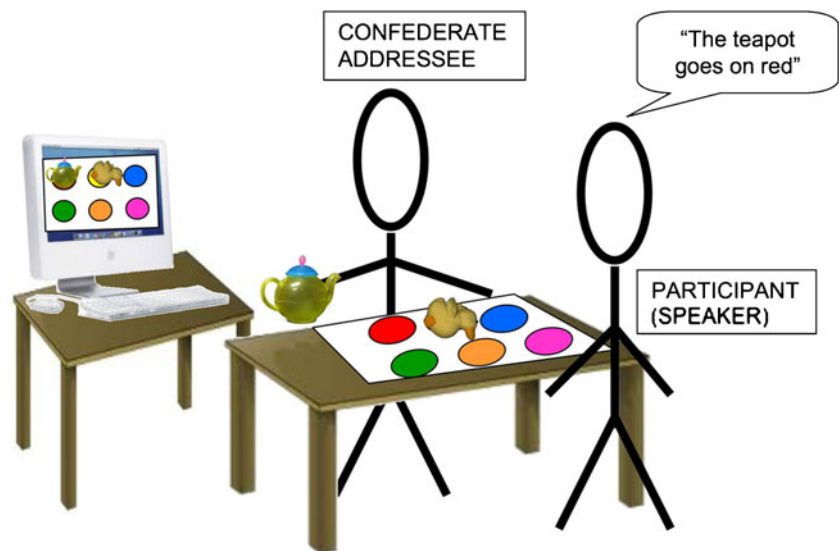
understanding of a target reference. We predicted that speakers would be most likely to attenuate pronunciations when they had immediate, salient information that their addressee already had understood. Imagine that Alex says to Elise, *I need to sign this. Can you hand me the pen?* Elise may predict his request and pick up the pen before he mentions it. If Alex believes that Elise's action reflects her understanding of his intended meaning, he may then attenuate the word *pen*.

A critical feature of this example is that the speaker and listener share access to all relevant aspects of the discourse. Of particular importance is the predictability of the reference to the pen, because predictability favors reduction (e.g., Bell et al., 2009; Jurafsky et al., 2001; Watson et al., 2008). Predictability is traditionally considered an informational property of the situation. We propose a variation of this view, whereby predictability effects are modulated by feedback from the addressee that demonstrates that he or she understands the predicted information. We tested this idea in a referential communication task, in which the participant gave instructions to a confederate addressee about how to move objects on a table (see Fig. 1). There were two objects per trial. We manipulated whether the addressee *anticipated*, by picking up the second object before it was mentioned, or *waited* for the instruction.

Our first goal was to examine pronunciations as a function of the addressee's behavior. If addressee feedback influences acoustic variation, pronunciations should reduce in the anticipation condition. Our second goal was to consider the potential mechanisms by which audience design might influence pronunciations. According to the *comprehension facilitation* view of audience design (e.g., Lindblom, 1990), speakers choose a pronunciation on the basis of assumptions about the addressee's knowledge. Similarly, the givenness hierarchy suggests that speakers select lexical items as a function of the addressee's attention (e.g., Gundel, Hedberg, & Zacharski, 1993). Variants of this mechanism depend on such questions as whether the speaker represents specific or general addressee knowledge (Brennan & Schober, 2001), but the key property of this view is that assumptions about the addressee's knowledge feed directly into decisions about articulation. This view predicts that acoustic variation will be specific to the anticipated information, as *pen* is in the example above.

Alternatively, addressee anticipation may lead to acoustic reduction via *production facilitation*, thus predicting feedback effects on acoustic prominence primarily in the planning regions of the utterance—that is, at utterance onset (Clark & Wasow, 1998). We hypothesized that anticipation might facilitate production in numerous ways—for example, by highlighting the predictability of an event, drawing attention to the target object, or sending a signal that the addressee is attentive—allowing the speaker to focus on message formulation. Even if the speaker is not actively monitoring the addressee, the anticipation action provides one piece of evidence that the

**Fig. 1** Diagram of experimental layout



addressee already knows which object to move, reducing the speaker's uncertainty. Indeed, evidence suggests that planning is facilitated when interlocutors are "on the same page," in that prosodic entrainment leads to shorter latencies to begin speaking (Levitan, Gravano, & Hirschberg, 2011).

One advantage of this task is that it was relatively natural. Therefore, we also expected speakers to attenuate their utterances in other ways for anticipating addressees—for example, by using fewer words. Analyses of lexical choices thus served as a test of the manipulation.

## Experiment 1

### Method

**Participants** A group of 17 undergraduates at the University of North Carolina (UNC) Chapel Hill participated for course credit. One was excluded due to technical problems.

**Procedure** In an interactive referential communication task, the participants gave instructions to a confederate to place objects on a board of six colored dots (Fig. 1; a video may be seen at [www.unc.edu/~jarnold/papers/AKPsupporting.html](http://www.unc.edu/~jarnold/papers/AKPsupporting.html)). Each trial had two instructions, and the target was the object in the second.

The participant stood on one side of the table holding the display board, facing the addressee. Behind the addressee, a computer screen, controlled by an experimenter, displayed each object's location. Participants were instructed to tell the confederate to put each object on the correct colored dot; the confederate could not see the computer screen, and the participant was not allowed to touch or point at the objects.

The objects for each trial were stacked in boxes next to the table. At trial onset, the addressee placed the objects for that

trial on the table. When the objects touched the table, the experimenter brought up a picture displaying the first object's location (e.g., duck on yellow), and the participant gave the first instruction. After the confederate had placed the object, the experimenter displayed the second, critical stimulus.

Stimulus onset was recorded with a beep, sent inaudibly from the computer via a sound mixer to a Marantz PMD670 digital recorder. The participant wore a Lavalier microphone connected to the sound mixer/recorder. A video camera recorded the confederate's actions.

**Design and materials** The 48 trials occurred in two blocks, each with a different confederate addressee. The critical manipulation was whether the addressee displayed knowledge of the predictability of the target object. The confederate in the *anticipating* condition picked up *both* objects (one in each hand) after hearing the first instruction, anticipating the use of the second object. In the *waiting* condition, the confederate simply waited to hear each instruction before moving the object. Both confederates behaved naturally and cooperatively; the behavioral contrast was consistent with natural variation between individuals.

Each participant saw 24 trials with each confederate; each trial included two objects. Thus, there were two sets of 24 pairs of objects. The targets (second objects) in each set were matched by syllable lengths, phoneme lengths, and log frequencies. We counterbalanced the order of the conditions across participants and by confederates, as well as the assignment of the four individual confederates to the conditions.

In order to encourage the speaker to attend to the addressee and the objects, we incorporated an "error" in the picture for the filler object on Trial 3. The experimenter apologized and suggested that the participant subsequently check the objects on the table.

We analyzed the effects of anticipation on both number of words and acoustic reduction, measured by the duration and intensity. We analyzed three regions of the participants' speech: (1) the latency to initiate an utterance, measured from picture onset; (2) duration of the determiner (*the*); and (3) duration of the target noun (e.g., *teapot*).

**Predictions** If audience design affects articulation, anticipation should promote shorter durations and/or lower intensity of the instruction. Under the comprehension facilitation view, such effects should be greatest on the target noun and should possibly extend to other regions. Alternatively, if the anticipation manipulation affected utterance planning, its effect should primarily emerge in the planning regions—that is, the latency to begin speaking and the determiner, which was usually the first or second word of the utterance in this data set.

## Results

From the 768 trials, 13 items were excluded (two for confederate error, 11 for object identification error or repair).

All results were analyzed with maximum-likelihood-based multilevel regressions in R. For each dependent variable, we constructed a model in three steps, fitting (1) random effects, (2) control predictors, and (3) the predictors of interest: trial order (with later trials predicted to be shorter) and condition (anticipating vs. waiting).

We modeled random effects (Table 1) for both a) subjects and b) either items (for the lexical analyses) or the target word used (for the acoustic analyses).<sup>1</sup> We first fit a “maximal” random-effects structure, including random intercepts and slopes for both predictors of interest, relative to subjects and items/targets. If any random-effects-only model failed to converge, or if any pair of random effects had a correlation greater than .9, we removed the least theoretically relevant effect and refit the model.

Control predictors (Table 2) were fit stepwise, starting with the most theoretically relevant. Predictors were excluded if the *t* value was 1.5 or less or if the predictor was collinear with a previously introduced predictor (bivariate correlation > .2).

For the predictors of interest, log-likelihood ratio testing was conducted, beginning with trial order, then condition, then the interaction between the two. Critical predictors were only retained if they increased the fit of the model. Only significant tests are reported below.

<sup>1</sup> On 89% of trials, speakers used the intended object word, but some objects evoked multiple names (e.g., cheetah/leopard/tiger/animal). The word was chosen over item for random effects because it heavily constrains duration.

**Lexical variation** The participants were free to phrase their instructions in any way that they wished, which led to substantial variation, from *and you can put the tiger on the orange, yup* to simply *tiger orange*, or even in a few cases dropping the object noun entirely (*and yellow*). This variation, indexed by the number of words used, varied significantly by condition: the target object was described with fewer words in the anticipating (mean 1.57) than in the waiting (mean 1.67) condition [ $\chi^2(1) = 6.06, p < .01; \beta = -.57, SE(\beta) = .2, t = -2.83$ ]. Fewer words were also used in later trials [ $\chi^2(1) = 11.22, p < .0008; \beta = -.025, SE(\beta) = .0044, t = -5.69$ ]. A similar pattern occurred in an analysis of words in the entire utterance (mean 5.0 in anticipating; mean 5.5 in waiting condition) [ $\chi^2(1) = 5.22, p < .02; \beta = -.11, SE(\beta) = .041, t = -2.79$ ; effect of trial order:  $\chi^2(1) = 25.68, p < .0001; \beta = -.0058, SE(\beta) = .0014, t = -3.99$ ]. In sum, anticipating addressees evoked less information than did waiting ones.

**Acoustic variation** All single-label references ( $N = 657$ ) were analyzed with Praat 5.2.01 (Boersma & Weenink, 2011), excluding trials with unusable recordings ( $n = 9$ ), multiword expressions ( $n = 50$ ), pronominal *that* ( $n = 2$ ), and zeros ( $n = 37$ ). The latency analysis excluded three outliers (>3 *SDs* above the mean), 20 trials in which the participant began speaking before the stimulus appeared, and one trial in which the beep was not recorded.

We analyzed the log durations of the latency to begin the utterance, the determiner, and the object word, along with the amplitude of the determiner and the object word (Table 3). We controlled for speaking rate, indexed by dividing the total time between the onset of the utterance and the offset of the color word by the number of words in that region.

Speakers began speaking more quickly with anticipating addressees [ $\chi^2 = 4.29, p < .03; \beta = -.038, SE(\beta) = .017, t = -2.22$ ] and for later trials [ $\chi^2 = 15.8, p < .0001; \beta = -.0042, SE(\beta) = .00084, t = -5.01$ ]. They also produced shorter determiners for anticipating addressees [ $\chi^2 = 5.1, p < .02; \beta = -.039, SE(\beta) = .017, t = -2.34$ ]. Critically, anticipating condition had no effect on object word duration [ $\chi^2 = 0.3, p > .58; \beta = -.003, SE(\beta) = .0063, t = -0.55$ ]; neither did condition affect amplitude of either the determiner or the object word.

## Discussion

Addressee anticipation had clear effects on the speaker's instructions, but only for some processes. Speakers used fewer words when describing targets to anticipating addressees, confirming that the manipulation had the expected effect. This finding also demonstrated that the manipulation did not simply signal speakers to initiate utterances more

**Table 1** Random effects in each model, shown in order of priority for including in the models

		Duration			Intensity		Lexical/Phrasal	
		Delay	Det.	Obj. Word	Det.	Obj. Word	# of Words in Sent.	# of Words in Target Phrase
Subject	Intercept	0.001	0.002	0.001	39.76	208.9	3.04	0.25
	Condition	< 0.001	%%	%%	%%	%%	0.91	0.008
	Trial	%%	%%	%%	0.003	0.229	%%	%%
	Condition * Trial	–	–	–	–	–	0.003	–
Item/word	Intercept	0.014	%%	0.002	0.73	0.851	0.04	0.001
	Condition	%%	0.003	%%	%%	%%	%%	%%
	Trial	< 0.001	%%	%%	< 0.001	%%	%%	%%
	Condition * Trial	–	–	%%	–	–	%%	–

Dashes indicate that the effect was not included due a lack of convergence, and percent signs indicate that the effect was not included due to a high correlation (absolute value > .9) with other random effects. Variance estimates for the final models are listed, otherwise.

quickly, in that anticipation had lexical effects throughout the utterance.

Critically, anticipation affected the acoustic prominence of determiners. To our knowledge, this is the first evidence that addressee feedback affects spoken word durations (cf. Bard et al., 2000; Bard & Aylett, 2004; Galati & Brennan, 2010; Smiljanic & Bradlow, 2009)

However, anticipation did not affect the pronunciation of the target word. This is inconsistent with the comprehension facilitation view of audience design, which predicts that if addressees show evidence of having predicted the target

reference, the target word itself should be reduced in prominence. Instead, anticipation led to shorter utterance initiation times and shorter determiners. Importantly, these reductions were not due to changes in speech rate, which was controlled in the model. Utterance onset regions, and particularly determiners, have been associated with utterance planning (Clark & Wasow, 1998). This suggests that addressee anticipation affected utterance planning.

Yet before we conclude that anticipation had no effect on object word reduction, we must consider whether speakers signaled prominence with other acoustic features than

**Table 2** Control variables in each model

		Duration			Intensity		Number of Words	
		Delay	Det. ( <i>the</i> )	Target Word (Head Noun)	Det.	Target Word	#Words in Sent.	#Words in Target Phrase
Rate effects	Time per syllable <sup>a</sup>	$t = 2.87$	$t = 13.19$	$t = 16.13$	–	$t = -1.87$		
Word-level effects	<i>The</i> intensity	–	$t = -3.49$	NA	NA	$t = 5.85$		
	Target intensity	–	NA	–	$t = 4.51$	NA		
	# of words in target area	–	NA	–	NA	NA		
Context effects	Syllables in target word	$t = -1.49$	–	$t = 10.91$	$t = -1.82$	–		
	Word before <i>the</i> <sup>b</sup>	NA	$t = -3.7$	NA	$t = -3$	–		
	Word after <i>the</i> /before target	NA	–	–	$t = -2.56$	–		
Sentence-level effects	Word after target	NA	NA	–	NA	–		
	# of words in sentence	–	–	–	–	–		
Confederate	Disfluency in target sentence	–	–	–	–	–	–	–
	Confederate ID	–	–	$t = 1.96$	$t = -2.64$	–	$t = 1.55$	–
Participant	Gender = female	$t = -2.33$	–	–	–	–	–	–
Item	Item set	–	–	–	–	–	–	–

Dashes indicate that the variable was not included in the model. The  $t$  values mark significant effects and the direction of the effect (positive/negative). <sup>a</sup> The number of syllables for each word was coded from the transcript. A subset of utterances for each speaker was compared with the audio files, and no discrepancies were found between the pronounced and transcribed syllables. <sup>b</sup> We categorized the preceding and following words into two groups: (1) nothing, a pause, or a disfluency (or, in the case of the word following the head, a “telegraphic speech” color word, as in *tiger orange*), or (2) other.

**Table 3** Average durations for critical regions, in milliseconds

	Latency	Determiner	Object Word
Anticipating	545	126	409
Waiting	586	141	412

duration and amplitude (Wagner & Watson, 2010). We tested this possibility in [Experiment 2](#), in which participants rated the intelligibility of the object words from [Experiment 1](#). We followed the method of Galati and Brennan (2010), who found audience-design effects with intelligibility ratings but not with durational analyses.

## Experiment 2

### Method

*Participants* A group of 16 different undergraduates at UNC Chapel Hill participated for course credit.

*Procedure, design, and materials* We selected 68 pairs of target words (136 tokens) from the [Experiment 1](#) responses. Participants listened to each pair, repeated the word, and rated each token for clarity, from 1 (*low clarity*) to 5 (*high clarity*). They were instructed to pace themselves and to play both sound files as often as needed.

We included as many participants and items as possible from both conditions, with the following criteria: Each pair contained a token of the same word from each condition (anticipating vs. waiting) and from nearly identical lexical and syntactic contexts. Object words in [Experiment 1](#) did not differ in duration across conditions, so pairs were matched on duration (average 27-ms difference). Each token in the pair came from a different speaker, given the design of [Experiment 1](#). Four pairs that did not meet these criteria were erroneously used but were excluded from the final analysis. Four lists counterbalanced item order and token order.

By closely matching tokens in each pair on duration and contextual factors, we hoped to encourage participants to focus their attention on other acoustic indicators of intelligibility, in order to maximize the chance of detecting a systematic effect of condition on intelligibility, broadly construed.

Spectral noise gating was performed on the original recordings to reduce background noise using Audacity software (Audacity Team, 2011). For each sound file, a noise profile was created by performing a Fourier analysis on a few seconds at the beginning of the recording and reducing those constituent frequencies by 24 dB throughout the rest of the file. Target words were extracted from the filtered

sound files, and their average intensity was normalized to 70 dB using Praat.

### Results and discussion

A total of six of the 2,048 ratings were excluded because the participant failed to repeat the word. The ratings were analyzed using the same model-building procedure as in [Experiment 1](#). The final model included random intercepts for subject, token, and speaker, and random slopes for the condition-by-speaker and condition-by-token interactions. The condition-by-subject slope was excluded due to high correlation with other effects.

Ratings did not differ between the anticipating (mean = 3.11) and waiting (mean = 3.14) conditions [ $\chi^2(1) = 0.95$ ,  $p > .33$ ;  $\beta = .14$ ,  $SE(\beta) = .14$ ,  $t = 1.0$ ]. Instead, higher ratings were predicted by the following control variables: (1) log word duration, (2) later trials, and (3) being the second token of the pair.

The rating measure tested whether the anticipation manipulation in [Experiment 1](#) led to systematic variation in intelligibility, even if this was expressed with different combinations of prosodic features across participants or trials. Any such combination would have led to perceptual differences on this task, yet no condition effects were detected. In sum, object words were no more intelligible for waiting than for anticipating addressees.

### General discussion

In this study, we used a novel method for testing the effects of audience design on acoustic prominence, focusing on addressee behavior. Two important results emerged. First, speakers did indeed adjust their pronunciation as a function of addressee feedback, shortening the word *the* when the addressee anticipated the referent. This supports the hypothesis that speakers use acoustically reduced tokens when the addressee has contextual support for understanding the speaker's message, and contrasts with claims that the processes underlying acoustic reduction work too quickly to take audience design into account (Bard & Aylett, 2004).

Second, the effects of addressee feedback were not consistent with the comprehension facilitation view of audience design. If they were, we should have seen reduction of the object word, because it encodes the relevant anticipated information. Yet there was no evidence that anticipation led to a reduction in object word duration, amplitude, or intelligibility.

The pattern of results instead showed that the addressee's behavior influenced the planning regions: that is, utterance initiation and determiner duration. This

suggests that addressee feedback may influence acoustic reduction indirectly, via production facilitation. In short, audience design and production processing are likely interrelated.

This idea is consistent with view that acoustic reduction results from fast, automatic processes, many of which are production-internal (Kahn & Arnold, 2011; Lam & Watson, 2010). This view does not exclude the possibility of reduction on the target word, which should depend on the degree of production facilitation in a given task. Thus, while our data contrast with Bard and Aylett's (2004) failure to find any effects of audience design on intelligibility, they are consistent with their theoretical framework. Our results raise questions about the relationship between planning, acoustic reduction, and other kinds of evidence about addressee knowledge and attention. In other tasks, such planning effects may underlie addressee effects on intelligibility (see, e.g., Galati & Brennan, 2010).

An open question is how quickly addressee feedback can impact acoustic decisions. In this study, feedback varied by block, so speakers could make general inferences about their addressee's behavior for the entire block. Given the opportunity, speakers may respond to dynamic feedback (Clark & Krych, 2004). Alternatively, acoustic adjustments may result only from global inferences (cf. Anderson, Bard, Sotillo, Newlands, & Doherty-Sneddon, 1997).

These findings alter an ongoing debate about the role of audience design in the production of acoustic prominence. The question is no longer whether speakers vary pronunciation on the basis of addressee needs—we know that they do. Our research instead raises questions about the kind of feedback that matters, and how audience behavior relates to known production-internal processes.

**Author note** This work was funded by National Science Foundation Grant BCS-0745627. We gratefully acknowledge the assistance of Kellen Carpenter, Alex Christodoulou, Elise Rosa, Renske Hoedemaker, Kayla Finch, and Molly Bergeson.

## References

- Anderson, A. H., Bard, E. G., Sotillo, C., Newlands, A., & Doherty-Sneddon, G. (1997). Limited visual control of the intelligibility of speech in face-to-face dialogue. *Perception & Psychophysics*, *59*, 580–592. doi:10.3758/BF03211866
- Audacity Team. (2011). Audacity (Version 1.3 beta) [Computer program]. Retrieved January 11, 2011, from <http://audacity.sourceforge.net>
- 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.
- Balota, D. A., Boland, J. E., & Shields, L. W. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory and Language*, *28*(1), 14–36.
- 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. doi: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. doi:10.1016/j.jml.2008.06.003
- Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer (Version 5.2.01) [Computer program]. Retrieved February 1, 2010, from [www.praat.org](http://www.praat.org)
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1482–1493. doi:10.1037/0278-7393.22.6.1482
- Brennan, S. E., & Schobor, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language*, *44*, 274–296. doi:10.1006/jmla.2000.2753
- Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, *50*, 62–81. doi:10.1016/j.jml.2003.08.004
- Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology*, *37*, 201–242.
- Ferreira, V. S., & Dell, G. S. (2000). The effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, *40*, 296–340.
- 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. doi:10.1016/j.jml.2009.09.002
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions. *Language*, *69*, 274–307.
- Halliday, M. A. K. (1967). Notes on transitivity and theme in English (Part 2). *Journal of Linguistics*, *3*, 199–244.
- Horton, W. S., & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, *59*, 91–117.
- Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee & P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam, The Netherlands: John Benjamins.
- Kahn, J., & Arnold, J. E. (2011). *Speakers reduce because of their own activated representations*. Paper presented at the CUNY Conference on Human Sentence Processing, Stanford, CA.
- Kahn, J., & Arnold, J. E. (2012). *A processing-centered look at the contribution of givenness to durational reduction*. Manuscript under review.
- Lam, T. Q., & Watson, D. G. (2010). Repetition is easy: Why repeated referents have reduced prominence. *Memory & Cognition*, *38*, 1137–1146. doi:10.3758/MC.38.8.1137
- Levitan, R., Gravano, A., & Hirschberg, J. (2011). Entrainment in speech preceding backchannels. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 113–117). Portland, OR: Association for Computational Linguistics.
- Lindblom, B. (1990). Explaining variation: A sketch of the H and H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech*

- production and speech modelling* (pp. 403–439). Dordrecht, The Netherlands: Kluwer.
- Smiljanic, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Linguistics and Language Compass*, 3, 236–264.
- Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language & Cognitive Processes*, 25, 905–945.
- Watson, D. G., Arnold, J. E., & Tanenhaus, M. K. (2008). Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 106, 1548–1557. doi:10.1016/j.cognition.2007.06.009
- Wright, R. A. (2004). Factors of lexical competition in vowel articulation. In J. J. Local, R. Ogden, & R. Temple (Eds.), *Laboratory phonology* (Vol. 6, pp. 26–50). Cambridge, U.K.: Cambridge University Press.