Repetition is easy: Why repeated referents have reduced prominence

Tuan Q. Lam (tuanlam2@illinois.edu)
Department of Psychology, 603 E. Daniel Street
Champaign, IL 61820 USA

Duane G. Watson (dgwatson@illinois.edu)
Department of Psychology, 603 E. Daniel Street
Champaign, IL 61820 USA

Abstract

The goal of this study was to determine whether effects of repetition and predictability on the duration of referring expressions are independent, and whether factors that affect the acoustic properties of a word are the result of speaker-centered processes or listener-centered processes. In two experiments using a referential communication task, we found that informational predictability and repetition independently affect the intensity (loudness) of referring expressions, but only repetition affects their duration. We conclude that the acoustic properties of a word are determined by both listener and speaker-centered processes.

Keywords: reference; prosody; prominence; production, audience design

Introduction

In this study we examine the factors that affect the acoustic realization of referring expressions. A great deal of work has shown that the more predictable a word is, the less prominent it is (Bell et. al, 2003; Bell et al., 2009; Fowler & Housum, 1987; Jurafsky, Bell, Gregory, & Raymond, 2001; Pluymaekers, Ernestus, & Baayen, 2005a; Watson, Arnold, & Tanenhaus, 2007). This tends to be reflected acoustically through lower intensity, lower F0, and shortened duration of the word. Although it is clear that predictability and prominence are linked, there is less agreement as to why this link exists. In this study, we investigate the relationship between acoustic prominence and informational redundancy, and examine whether effects of redundancy on acoustic prominence are listener-driven or whether they are linked to speaker-centered production processes.

In explaining the link between predictability and prominence, some researchers have taken information theoretic approaches and have argued that prominence differences are the result of speakers attempting to facilitate comprehension (Aylett & Turk, 2004; Frank & Jaeger, 2008). Under this view, speakers attempt to produce a signal in which the probability of an element being recognized by the listener remains relatively constant throughout an utterance. They propose that prosodic prominence's primary role is to smooth the information content of the signal via reduction in syllable duration when a word is redundant: expected words are produced with shorter durations and unexpected words are produced with longer durations so that the amount of information is evenly distributed over time, which in turns, makes the word easier to detect.

On the other hand, the link between predictability and acoustic prominence may be a result of a speaker-centered process (see Bell et al. 2009, for a review). Bell et al. (2009) propose that the speed of lexical retrieval is linked to articulatory planning: if lexical retrieval is difficult, then speakers will spend more time producing the word. According to this model, context will make some words more accessible than other words. These words in turn will be easier to retrieve and therefore, less prominent. Thus, repeated referents are easier to retrieve both because of prior context, and because of residual activation from previous productions, and this leads to reduction.

One way to test whether reduction is the result of information theoretic principles or the result of processes related to speech production is to test whether effects of repetition and predictability are independent. Previous studies on repetition and predictability have relied primarily on corpus data (e.g. Aylett & Turk, 2004; Bell et al., 2009). In natural speech, repetition and predictability are highly correlated: repeated words are more predictable than non-repeated words (Arnold, 1998). Thus, it is difficult to know whether effects of repetition are separate from effects of predictability using natural speech. In Experiments 1a and 1b, we address this question by looking for independent effects of repetition and speaker expectation. Contexts were created in which repeating a word was unexpected, reversing the correlation that is typically seen in natural speech (Experiment 1a). As a control, contexts were also created in which repeating a word and producing a new word was equally likely (Experiment 1b).

If repeated reference causes reduction even when the repeated word is less predictable, then this would suggest that reduction is linked to factors in production. Previous activation of the word should facilitate production and lead to reduction, even if the word is not predicted. If repeated, less predictable target words are produced with longer duration than non-repeated, predictable words, this would support a redundancy avoidance account: predictable words are reduced to facilitate robust communication for the listener, independent of whether they have been produced before.

Experiment 1

Experiment 1 employed a referential communication task. There were two roles, director and matcher. The director's task was to describe sequences of two events occurring on his computer screen to the matcher who had to match the events on her screen. On each trial, participants were first shown a pair of objects on a computer screen. Then, one of the objects shrank. Following the shrinking event, either the same object would flash or the other object would flash.

There were two conditions, repeated mention and non-repeated mention. On repeated mention trials, an object shrank and then the same object flashed. On non-repeated mention trials, an object shrank and then the other object flashed. Typical utterances for both trial types are presented in (1) and (2).

Example:
(1) Repeated noun
The kite is shrinking…The kite is flashing.
(2)Non-repeated noun
The snail is shrinking…The kite is flashing.

There were two parts to the experiment: a training block and a test block. The training block was designed to manipulate the relationship between predictability and repetition. In Experiment 1a, a repeated mention event was much less likely than a non-repeated mention event, while in the Experiment 1b, repeated mention and non-repeated mentions were equally likely. The training block was followed immediately by a test block. For the test block, repeated mention trials and non-repeated mention trials were equally likely in both versions of the experiment. The critical target word was the production of the referent in the second event. The intensity, duration, and F0 of this word were measured.

Method

Participants Sixty-five undergraduate students from the University of Illinois at Urbana-Champaign participated in the experiment in exchange for either pay or course credit. Thirty-three students participated in Experiment 1a and thirty-two students participated in Experiment 1b.

One participant's data was excluded because the participant failed to produce both utterances. Two participants were excluded because of excessive background noise. Four participants' data were excluded because they produced pronouns instead of the name of the target object (in repeated trials).

Materials Twelve colored images were taken from a set of pictures used by Rossion and Pourtois (2001). These images are a colored version of the Snodgrass and Vanderwart (1980) pictures. The 12 images were grouped into six pairs for the experiment. Pairs were chosen so as to avoid semantic and phonetic overlap. One of the six image pairs appeared on each trial with one image on the left side and the other image on the right side.

For each trial, there were two events, a shrinking event and a flashing event. In the shrinking event, a target image shrank, and in the flashing event, a target image flashed. There were two conditions, a repeated noun condition and a non-repeated noun condition. In repeated noun trials, one of the two images shrank, and then the same image also flashed. In non-repeated noun trials, one of the images shrank and then the other image flashed. An example of a trial in the non-repeated condition is shown in Figure 1.

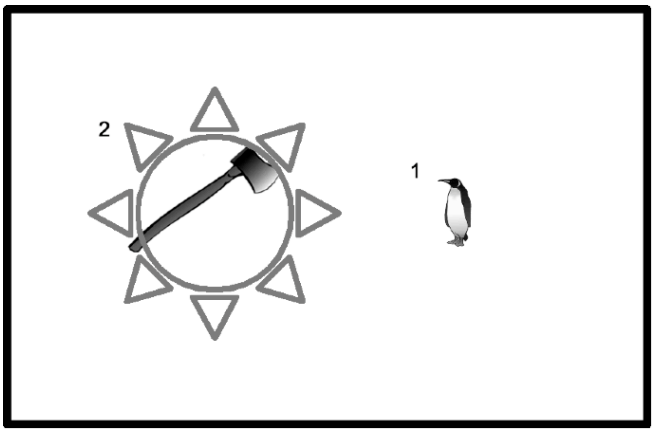

Figure 1: An example of a non-repeated trial in Experiment 1a and 1b.

There were a total of 108 trials, and each image pair appeared exactly 18 times, once in each left/right ordering, over the course of the experiment. The first 96 trials served as a training block to train participant on the event frequencies of the experiment. The frequency of the events varied across the training block in Experiments 1a and 1b. In Experiment 1a, six of the 96 training block trials were repeated noun trials. The remaining 90 trials were non-repeated noun trials. The conditions were pseudo-randomly permuted such that no repeated noun trial appeared within the first eight trials of Experiment 1a and there were never two back-to-back repeated noun trials during the training block. This manipulation created an expectation that repeated noun trials would be relatively unlikely. In Experiment 1b, exactly 48 trials were repeated noun trials, and 48 trials were non-repeated noun trials during the training block. As in Experiment 1a, the trial conditions were randomly permuted, only this time there were no controls on repeated noun trials. For both Experiment 1a and 1b, the final 12 trials served as a test block. For both Experiment 1a and 1b, half of the test block trials were repeated noun trials and half were non-repeated noun trials.

Procedure Before the experiment began, both participants were shown a video of two people completing the task. This video was meant to familiarize participants with the task and to prime the use of the target structure for each utterance.

After watching the video, participants were told that they would be playing the role of the director while the experimenter played the role of the matcher. The director's

role was to describe a pair of events that occurred on his screen. Meanwhile, the matcher's task was to manipulate the objects on her screen based on the director's description. The director and matcher sat at separate computers and were unable to see each other's screens. The pair completed six practice trials before beginning the actual experiment.

At the start of each trial, two images appeared on both screens. After 500 ms, one of the images on the director's screen began to shrink. The participant was instructed to begin speaking as soon as he knew what was happening during the first event. Upon hearing the director's utterance, the matcher clicked on a button to make the same object shrink on her screen. Exactly two seconds after the shrinking event, one of the images began flashing. Then the director began describing the second event. Again, the matcher matched the event on her screen. Then the trial ended.

Participants completed 108 trials. There was no break to indicate the end of the training block and the beginning of the test block. Participants' utterances were recorded during the test block. The intensity, duration, and F0 were measured for critical words.

Results and Discussion

The data were analyzed using a linear mixed effects regression model with subjects and items as random effects (Baayen, 2008). All reported p-values were obtained from Markov chain Monte Carlo sampling.

In both Experiments 1a and 1b, there were reliable effects of repetition such that repeated words were reduced. Predictability did not influence duration. In Experiment 1a, non-repeated words were produced with greater duration ($p<0.01$) than repeated words; however, there was no reliable difference for intensity. There were also no reliable differences in pitch across conditions. In Experiment 1b, non-repeated words were produced with greater duration ($p<0.001$) and intensity ($p<0.01$) than repeated words. As in Experiment 1a, there were no reliable differences in pitch across conditions for Experiment 1b.

There was a marginal interaction between condition and experiment for intensity ($p=0.07$). In Experiment 1a, non-repeated words and repeated words show no difference in intensity, while in Experiment 1b, non-repeated words were produced with more amplitude than repeated words.

Table 1: Mean duration (ms) and intensity (dB) across conditions for the target word in Experiment 1a and 1b. Standard errors are in parentheses.

Target Word	Duration	Intensity
Expt 1a Repeated	376 (14.3)	78.6 (1.08)
Expt. 1a Non-Repeated	393 (13.8)	78.7 (1.14)
Expt. 1b Repeated	339 (10.7)	77.0 (1.00)
Expt. 1b Non-Repeated	356 (10.3)	77.6 (0.97)

Repeated words were less prominent than non-repeated words even though they were less likely to occur, suggesting that production centered processes played a primary role in determining duration in this study. These results are inconsistent with information theoretic accounts.

Note however that although the duration results support production centered accounts, there were some differences in intensity across experiments. The pattern of results suggests that the lack of intensity differences observed in Experiment 1a was driven by the predictability manipulation. One possibility is that in Experiment 1a, predictability and repetition might have both had effects on intensity, but effects of predictability were not detectable, because predictability and repetition were not independently manipulated. Because repeated words were unexpected in Experiment 1a, it is impossible to determine whether the effects were from repetition, predictability, or some combination of these two factors.

Experiment 2

In Experiment 2, both repetition and predictability were independently manipulated in order to determine whether predictability and repetition had independent effects.

Participants were presented with a three by four array of 12 objects. As in Experiments 1a and 1b, there were two events on each trial. On every trial, one of the objects shrank and one of the objects flashed.

Unlike Experiments 1a and 1b, Experiment 2 used an explicit predictability cue. Predictability was manipulated by including the appearance of a circle around the expected referent of the second utterance immediately after the shrinking event but before the flashing event. On predictable trials, the circled object flashed. On unpredictable trials, the circled object was not the object that flashed. Repetition was also manipulated: either the same object shrank and then flashed or one object shrank and then a different object flashed. Predictability and repetition were crossed yielding four conditions: repeated expected, repeated unexpected, non-repeated expected, and non-repeated unexpected. As in Experiment 2, the critical word was the target in the second utterance.

Method

Participants Forty-five undergraduate students from the University of Illinois at Urbana-Champaign participated in this experiment in exchange for credit in an introductory psychology course. All participants were native speakers of American English.

Four participants' data were excluded from analysis due to interference from outside sounds. One participant was excluded from analysis for failing to produce the second utterance.

Materials The experiment used 96 color images from Rossion and Pourtois (2001). In each trial, 12 of these images were display in a three by four array. As in

Experiment 1a and 1b, the experiment was conducted using MATLAB with the Psychophysics toolbox installed.

As in Experiment 1a and 1b, there were two critical events during each trial, first a shrinking event and then a flashing event. Unlike in Experiment 1a and 1b, Experiment 2 used an explicit cue to manipulated predictability. After the shrinking event, a circle cued the participant as to which object would most likely flash during the flashing event. Exactly $11/12^{th}$ of the time, this object was indeed the object that flashed. On predictable trials, the object that was circled was also the object that flashed. On less predictable trials, the object that was circled was not the object that flashed. Repeated mention was manipulated by manipulating whether the same object or different objects shrank and flashed. Repeated targets occurred on exactly $1/12^{th}$ trials.

In this experiment, predictability and repeated mention were manipulated in a 2 x 2 factorial design leading to four conditions, repeated expected, repeated unexpected, non-repeated expected, and non-repeated unexpected.

At the beginning of a typical trial, 12 images appeared on the screen. After one second, one of the objects shrank. One second after the shrinking event, one of the objects was circled for 500ms and then disappeared. Exactly 500ms after the circle disappeared, one of the images flashed. An illustration of the non-repeated, unexpected condition is shown in Figure 2.

There were six critical trials in each condition for a total of 24 critical trials and 120 filler trials. For critical trials, the shrinking object, the circled object, and the flashing object had never previously been a target earlier in the experiment. The referent of the second event was the critical target word.

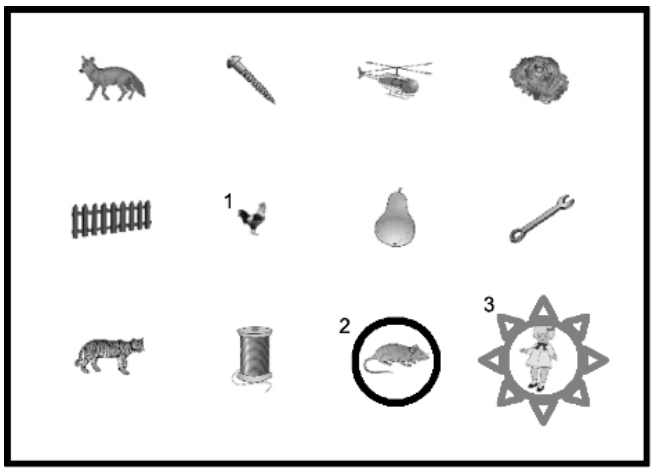

Figure 2: Example trial from Experiment 2, in the non-repeated, unexpected condition. First the rooster shrinks. Then the mouse gets circled. Finally, the doll flashes.

Procedure Before beginning the experiment, participants were shown a video of a participant completing this task. While watching the video, they were told that the circle would typically cue the shrinking event, but that sometimes, it would be an unreliable cue. They were instructed to only describe the shrinking event and the flashing event. Unlike Experiment 1a and 1b, there was no matcher.

After watching the video, participants completed eight practice trials with the experimenter present. Five of these trials were from the non-repeated expected condition. The remaining three trials were from each of the other three conditions.

At the beginning of a trial, 12 images appeared on the screen. Then one of the objects shrank. The participant then began describing this event. After the shrinking event, one of the objects was circled. Then one of the objects flashed. The participant also described the flashing event. When finished describing the flashing event, the participant pressed a key to move onto the next trial. The participant completed 144 trials with no breaks. On critical trials, the target word was analyzed for intensity, duration, and F0.

Results and Discussion

The data were analyzed using linear mixed effects regression with subjects and items as random factors. Non-repeated words had greater intensity than repeated words ($p<0.01$), and unexpected words had greater intensity than expected words ($p<0.001$). Non-repeated words were longer than repeated words ($p<0.0001$).

Table 2: Mean duration (ms) and intensity (dB) of target word across conditions in Experiment 2. Standard errors are in parentheses.

Target Word	Duration	Intensity
Repeated, Expected	421 (10.9)	58.2 (.828)
Non-repeated, Expected	446 (13.5)	58.4 (.814)
Repeated, Unexpected	428 (12.1)	58.7 (.906)
Non-repeated, Unexpected	458 (11.9)	59.2 (.874)

The results from Experiment 2 suggest that repetition and predictability affect prosodic prominence independently. Predictable words are produced with less intensity than unpredictable words. Repeated words are produced with less intensity and shorter duration than unrepeated words. The intensity results also suggest that the effect of repeated mention may have obscured the effect of predictability on intensity in Experiment 1a.

General Discussion

In this set of experiments, repeated mention led to reduced durations for target words. This was true both when repeated mention was predictable (Experiment 1b) and when it was less predictable (Experiment 1b). In Experiment 2 where predictability and repetition were manipulated independently, repeated mentions led to shorter duration for target words, while predictability did not affect duration.

The duration results are inconsistent with an information theoretic account of referring. These accounts argue that reduction is a result of speakers altering duration of words based on their predictability, but we found no link between predictability and duration here. Instead, these results are consistent with a speaker-centered account of reduction, in which reduction is partially determined by residual activation of previous productions of the word.

The intensity results differed from the duration results. There were no differences in intensity between repeated, unexpected words and non-repeated, expected words in Experiment 1a, but in Experiment 1b, where repeated words were more expected, repeated words were more reduced than non-repeated words. When repetition and predictability were manipulated separately as in Experiment 2, both predictability and repeated mention were shown to affect word intensity. Less predictable words were produced with more intensity than predictable words. Similarly, non-repeated words were produced with more intensity than repeated words.

Although the results here are inconsistent with information theoretic accounts of prominence, which propose that duration plays a primary role in smoothing the information signal, these results are not inconsistent with listener-centered accounts more generally. It is possible that speakers alter the intensity of their referring expressions for their listeners, but do not alter duration. In fact, intensity might be a more reliable cue for listeners as it can be detected relatively early in a word, while changes in duration may take more time for listeners to compute.

Previous work has shown a link between reduction and predictability, but the results here suggest that predictability does not influence duration as predicted by information theoretic accounts. Predictability is linked to changes in intensity. Duration was influenced by repetition, providing support for production-centered accounts of reduction. Overall, both predictability and repetition influence acoustic prominence of referents, but the data from these experiments suggest that they do so in different ways.

Acknowledgments

These experiments were funded by NIH grant: R01DC008774.

References

Arnold, J. E. (1998). *Reference form and discourse patterns*. Unpublished doctoral dissertation, Stanford University.

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.

Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge University Press.

Bell, A., Brenier, J., Gregory, M., Girand, C. and Jurafsky, D. (2009). Predictability Effects on Durations of Content and Function Words in Conversational English. *Journal of Memory and Language* 60:1, 92-111.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand C., Gregory, M. L., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of Acoustical Society of America*, 113 (2), 1001-1024.

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.

Frank, A. & Jaeger, T.F. (2008). Speaking Rationally: Uniform Information Density as an Optimal Strategy for Language Production. *The 30th Annual Meeting of the Cognitive Science Society (CogSci08)*. Washington, D.C. July, 2008.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In Bybee, J., & Hopper, P. (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam: Benjamins.

Rossion, B. & Pourtois, G. (2001). Revisiting Snodgrass and Vanderwart's object database: Color and texture improve object recognition. *Journal of Vision*, 1(3), 413a.

Snodgrass, J. G. & M. Vanderwart (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, 6(2), 174-215.

Watson, D.G., Arnold, J.E. & Tanenhaus, M.K. (2007). Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition*, 106, 1548-1557.