Toward a Unified Model of Redundancy Avoidance and Strategic Lengthening

T. Florian Jaeger & Celeste Kidd

Department of Brain & Cognitive Sciences, University of Rochester

{*fjaeger@bcs.rochester.edu; ckidd@bcs.rochester.edu*}

Reduction and Redundancy

Words that are more predictable in context are more commonly reduced [1, 2, 14].

However, content and function words appear to demonstrate an asymmetric pattern. Content words are reduced when predictable given the previous word, but function words are reduced when predictable given the *following* word. ***We present an account that unifies work on redundancy*** [1, 2, 13] ***with work on strategic lengthening*** [7]. We find that the apparent backward-predictability effect on function word reduction may be an artifact caused by speakers' tendency to slow production when the next word is unavailable.

Data

Corpus ***Paraphrase Stanford-Edinburgh LINK SWITCHBOARD*** [3, 4, 6, 10]

- 642 two-person telephone conversations between 358 speakers of American English.
- Word alignments created by automatic segmentation; a portion were reviewed manually.

Data set All occurrences of ***indefinite determiner "a"*** were extracted using TGrep2 [15].

- Indefinite article chosen because of the large number of instances (18,452) available.
- Disfluent cases, and cases with missing or unreliable variable information, were excluded, leaving 13,937 occurrences.
- Cases where the support for the relevant two- and three-word collocations (i.e. bigrams and trigrams) needed to calculate probability estimates occurred less than four times were also excluded, leaving ***5,317 occurrences***.

Data composition

	True	*False*
Vowel preceding	844	4,473
Vowel following	48	5,269
Disfluency preceding	332	4,985
Disfluency following	69	5,248
Pause preceding	90	5,227
Pause following	35	5,282

- For highly correlated factors, one factor was regressed against the other in a linear model and the residuals from the fit were used as variables in the mixed effects model.

Linear Mixed Model of "a" Duration

Control variables in the model

Speech rate Speakers produce words faster and with less articulatory detail during fast speech [2].

Gender Men speak faster and exhibit a higher degree of reduction in speech than females [2, 12, 17].

Age Older speakers tend to speak more slowly than younger ones [17].

Redundancy-based reduction variables in the model Words that are highly predictable in their context are more commonly reduced (pronounced shorter and with less articulatory detail) [1, 2, 14]. These words are redundant in that they provide less information about the message of the utterance. We include two measures of redundancy:

PRECEDING "a" ***FOLLOWING***
Chaos ***is a friend*** of mine.
Forward-predictability measure: log p(*a* | *is*)
Backward-predictability measure: log p(*a* | *friend*)

Log forward-predictability of "a" (bigram) The conditional probability of each determiner given the word preceding it.

Log backward-predictability of "a" (bigram) The conditional probability of each determiner given the word *following* it.

Availability-based variables in the model Previous studies show speakers structure utterances to provide themselves more time to plan difficult words and phrases (e.g., strategic lengthening [8], insertion of filled pauses [5], or insertion of optional function words [7]). However, current models of phonetic reduction do not consider the possible availability-based effects. We include three measures of the availability of the upcoming material:

NP length in words More complex NPs are harder to plan and may hence lead to strategic lengthening of the determiner introducing them.

Log frequency of following word Low-frequency words are harder to retrieve [13], causing speakers to extend the duration of a word to provide themselves more time for lexical access. This could cause speakers to lengthen "a" before low-frequency words.

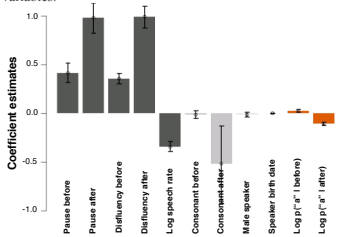
Log frequency measure: log p(*friend*)

Log trigram-predictability of following word The predictability of a word given its context affects its availability [9, 11] and hence should contribute to strategic lengthening.

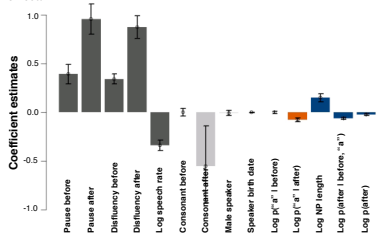
Trigram-predictability measure: log p(*friend* | *is a*)

Results and Discussion

Replication of previous results As expected [2, 14], only backward-predictability of "a"—but not forward-predictability —reached significance in a model without the availability-based variables.



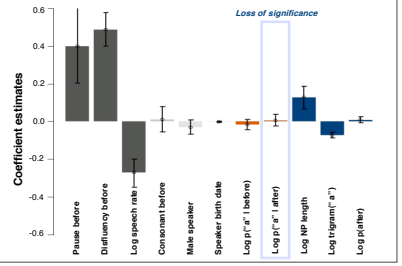
Effect of strategic lengthening All availability measures (in blue below) contribute significantly. Availability-trigram improved the model and reduced the backward-predictability effect.



Is the backward-predictability effect an artifact?

Collocations Further analysis revealed that the backward-predictability effect may be driven by collocations. "a" is often the first element of highly frequent collocations ("a lot of"). As would be expected if frequent collocations are retrieved together, "a" is pronounced significantly shorter in these contexts.

- Over 39% of the words following "a" co-occur with "a" more than 50% of all times (e.g. 90% of all occurrence of "lot" are preceded by "a"). This is the case for less than 0.2% of the words *preceding* "a".
- ***Backward-predictability of "a" is on average over 17-times higher than forward-predictability!***
- When highly frequent collocations (> 50% co-occurrence) are excluded from the data set, the backward-predictability becomes insignificant, but the trigram-availability and NP length remain clearly significant (same exclusion criteria applied to all probability measures).



Underlying Causes of Reduction

Accounting for duration variation in determiner "a" requires a unified probabilistic model that attributes reduction to three underlying causes: ***redundancy avoidance*** (predictable words are reduced), ***strategic lengthening*** before unavailable material, and joint storage and retrieval of ***collocations***.

Future work Extensions of the model will seek to test the following two hypotheses:

- ***Redundancy avoidance affects function words less than content words***, with the absence of frequency effects on function word reduction [2].
- ***Strategic lengthening affects function words more***: their high availability and occurrence before low-probability words (lexical heads) makes them ideal candidates for strategic lengthening.

References

[1] 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(1)*, 31-56.
[2] Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position of word form variation in English conversation. *Journal of the Acoustical Society of America, 113(2)*, 1001-1024.
[3] Bresnan, J., Carletta, J., Crouch, R., Nissim, M., Stoedtman, M., Wawer, T., & Zaenen, A. (2002). Paraphrase analysis for improved generation. In *LINK project: HRCR Edinburgh-CLSI Stanford*.
[4] Calhoun, S., Nissim, M., Steedman, M., & Brenier, J. (2005). A framework for annotating information structure in discourse. In *Proceedings of Frontiers in Corpus Annotation II: Pie in the Sky, ACL 2005 Conference Workshop*, 45-52. Ann Arbor, MI.
[5] Clark, H., & Fox Tree, J. (2002). Using uh and um in spontaneous speaking. *Cognition, 84*, 73-111.
[6] Deshmukh, N., Ganapathiraju, A., Gleeson, A., Hamaker, J., & Picone, J. (1998.) Resegmentation of Switchboard. In *Proceedings of the International Conference on Spoken Language Processing*, 1543-1546. Sydney, Australia.
[7] Ferreira, V. & Dell, G. (2000). Effect of Ambiguity and Lexical Availability on Syntactic and Lexical Production. *Cognitive Psychology, 40*, 296-340.
[8] Fox Tree, J. & Clark, H. (1997). Pronouncing "the" as "thee" to signal problems in speaking. *Cognition, 62*, 151-167.
[9] Frank, A. & Jaeger, T. F. (2008). Models of Speaker Choice in Production: Integrating Information and Availability. The 21st CUNY Sentence Processing Conference, Chapel Hill, NC.
[10] Godfrey, J., Holliman, E., & McDaniel, J. (1992). Switchboard: Telephone speech corpus for research and development. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 1*, 517-520. San Francisco, CA.
[11] Jaeger, T. F. (2006). Redundancy and syntactic reduction in spontaneous speech. PhD dissertation, Stanford University.
[12] Jaeger, T. F., Snider, N., Staum, L., & Jurafsky, D. (2006). (In)dependence of Lexical and Syntactic Production: that-reduction and omission in spontaneous speech. The 19th CUNY Sentence Processing Conference, New York, NY.
[13] Jescheniak, J. & Levelt, W. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 824-843.
[14] Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica, 62*, 146-159.
[15] Roland, D., Elman, J., & Ferreira, V. (2006). Why is that? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition, 98*, 245-272.
[16] Rohde, D. (2001). TGrep2 user manual.
[17] Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an Integrated Understanding of Speaking Rate in Conversation. In *Proceedings of Interspeech 2006*, 541-544. Pittsburg, PA.