Proceedings of the National Academy of Sciences of the United States of America 2018 vol.115 N22, pages 5720-5725

Nouns slow down speech across structurally and culturally diverse languages

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

© 2018 National Academy of Sciences. All rights reserved. By force of nature, every bit of spoken language is produced at a particular speed. However, this speed is not constant—speakers regularly speed up and slow down. Variation in speech rate is influenced by a complex combination of factors, including the frequency and predictability of words, their information status, and their position within an utterance. Here, we use speech rate as an index of word-planning effort and focus on the time window during which speakers prepare the production of words from the two major lexical classes, nouns and verbs. We show that, when naturalistic speech is sampled from languages all over the world, there is a robust cross-linguistic tendency for slower speech before nouns compared with verbs, both in terms of slower articulation and more pauses. We attribute this slowdown effect to the increased amount of planning that nouns require compared with verbs. Unlike verbs, nouns can typically only be used when they represent new or unexpected information; otherwise, they have to be replaced by pronouns or be omitted. These conditions on noun use appear to outweigh potential advantages stemming from differences in internal complexity between nouns and verbs. Our findings suggest that, beneath the staggering diversity of grammatical structures and cultural settings, there are robust universals of language processing that are intimately tied to how speakers manage referential information when they communicate with one another.

http://dx.doi.org/10.1073/pnas.1800708115

Keywords

Language processing, Language universals, Nouns, Speech rate, Word planning

References

[1] de Saussure F (1916) Cours de linguistique générale (Payot, Lausanne, Switzerland).

[2] Blasi DE, Wichmann S, Hammarström H, Stadler PF, Christiansen MH (2016) Sound-meaning association biases evidenced across thousands of languages. Proc Natl Acad Sci USA 113:10818–10823.

[3] Vaissière J (1983) Language-independent prosodic features. Prosody: Models and Measurements, Springer Series in Language and Communication, eds Cutler A, Ladd DR (Springer, Heidelberg), pp 53–66.

[4] Bybee JL (2010) Language, Usage and Cognition (Cambridge Univ Press, Cambridge, UK).

[5] Stivers T, et al. (2009) Universals and cultural variation in turn-taking in conversation. Proc Natl Acad Sci USA 106:10587–10592.

[6] Dellwo V, Leemann A, Kolly M-J (2015) Rhythmic variability between speakers: Articulatory, prosodic, and linguistic factors. J Acoust Soc Am 137:1513–1528.

[7] Ordin M, Polyanskaya L, Laka I, Nespor M (2017) Cross-linguistic differences in the use of durational cues for the segmentation of a novel language. Mem Cognit 45: 863–876.

[8] Gahl S (2008) Time and thyme are not homophones: The effect of lemma frequency on word durations in spontaneous speech. Language 84:474–496.

[9] Bard EG, et al. (2000) Controlling the intelligibility of referring expressions in dialogue. J Mem Lang 42:1–22.

[10] Aylett M, Turk A (2004) The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. Lang Speech 47:31–56.

[11] Jaeger TF (2010) Redundancy and reduction: Speakers manage syntactic information density. Cognit Psychol 61:23–62.

[12] Bell A, Brenier JM, Gregory M, Girand C, Jurafsky D (2009) Predictability effects on durations of content and function words in conversational English. J Mem Lang 60: 92–111.

[13] Fox Tree JE, Clark HH (1997) Pronouncing “the” as “thee” to signal problems in speaking. Cognition 62:151–167.

[14] de Jong NH (2016) Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. Int Rev Appl Linguist Lang Teach 54:113–132.

[15] Bates E, et al. (2003) Timed picture naming in seven languages. Psychon Bull Rev 10: 344–380.

[16] Jescheniak JD, Levelt WJM (1994) Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. J Exp Psychol Learn Mem Cogn 20:824–843.

[17] Levelt WJ, Roelofs A, Meyer AS (1999) A theory of lexical access in speech production. Behav Brain Sci 22:1–38, discussion 38–75.

[18] Vigliocco G, Vinson DP, Druks J, Barber H, Cappa SF (2011) Nouns and verbs in the brain: A review of behavioural, electrophysiological, neuropsychological and imaging studies. Neurosci Biobehav Rev 35:407–426.

[19] Szekely A, et al. (2005) Timed action and object naming. Cortex 41:7–25.

[20] Jaeger TF, Buz E (2017) Signal reduction and linguistic encoding. The Handbook of Psycholinguistics, eds Fernández EM, Cairns HS (John Wiley & Sons, Hoboken, NJ), pp 38–81.

[21] Conwell E (2017) Prosodic disambiguation of noun/verb homophones in child-directed speech. J Child Lang 44:734–751.

[22] Kibrik AA, Khudyakova MV, Dobrov GB, Linnik A, Zalmanov DA (2016) Referential choice: Predictability and its limits. Front Psychol 7:1429.

[23] Arnold JE (2008) Reference production: Production-internal and addressee-oriented processes. Lang Cogn Process 23:495–527.

[24] Ariel M (2014) Accessing Noun-Phrase Antecedents (Routledge, London).

[25] Gatt A, Krahmer E, van Deemter K, van Gompel RPG (2014) Models and empirical data for the production of referring expressions. Lang Cogn Neurosci 29:899–911.

[26] Levinson SC (2000) Presumptive Meanings. The Theory of Generalized Conversational Implicature (MIT Press, Cambridge, MA).

[27] Grosz BJ, Joshi AK, Weinstein S (1995) Centering: A framework for modeling the local coherence of discourse. Comput Linguist 21:202–225.

[28] Bickel B (2003) Referential density in discourse and syntactic typology. Language 79: 708–736.

[29] Stoll S, Bickel B (2009) How deep are differences in referential density? Crosslinguistic Approaches to the Psychology of Language: Research in the Tradition of Dan Isaac Slobin, eds Guo J, et al. (Psychology Press, New York), pp 543–555.

[30] Fox B (1987) Discourse Structure and Anaphora: Written and Conversational English (Cambridge Univ Press, Cambridge, UK).

[31] Givón T (1983) Topic continuity in discourse: An introduction. Topic Continuity in Discourse: A Quantitative Cross-Language Study, ed Givón T (John Benjamins, Amsterdam, The Netherlands), pp 1–41.

[32] Chafe WL (1994) Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Experience in Speaking and Writing (Univ of Chicago Press, Chicago).

[33] Seifart F (2012) The threefold potential of language documentation. Potentials of Language Documentation: Methods, Analyses, and Utilization, Language Documentation & Conservation Special Publication, eds Seifart F, et al. (Univ of Hawai‘i Press, Manoa, HI), pp 1–6.

[34] Kemmerer D (2014) Word classes in the brain: Implications of linguistic typology for cognitive neuroscience. Cortex 58:27–51.

[35] Indefrey P, Levelt WJM (2004) The spatial and temporal signatures of word production components. Cognition 92:101–144.

[36] Bosker HR, Pinget A-F, Quené H, Sanders T, de Jong NH (2013) What makes speech sound fluent? The contributions of pauses, speed and repairs. Lang Test 30:159–175.

[37] Koreman J (2006) Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech. J Acoust Soc Am 119:582–596.

[38] Seyfarth S (2014) Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. Cognition 133:140–155.

[39] Sóskuthy M, Hay J (2017) Changing word usage predicts changing word durations in New Zealand English. Cognition 166:298–313.

[40] Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? Behav Brain Sci 33:61–83, discussion 83–135.

[41] Levinson SC (2016) Turn-taking in human communication–Origins and implications for language processing. Trends Cogn Sci 20:6–14.

[42] Levelt WJM (1989) Speaking: From Intention to Articulation (MIT Press, Cambridge, MA).

[43] Hagoort P, Levinson SC (2014) Neuropragmatics. The Cognitive Neurosciences, eds Gazzaniga MS, Mangun GR (MIT Press, Cambridge, MA), pp 667–674.

[44] Moran S, et al. (2018) A universal cue for grammatical categories in the input to children: Frequent frames. Cognition 175:131–140.

[45] Anand P, Chung S, Wagers M (2015) Widening the net: Challenges for gathering linguistic data in the digital age. Available at https://www.nsf.gov/sbe/sbe_2020/2020_pdfs/Wagers_Matthew_121.pdf. Accessed January 25, 2018.

[46] Norcliffe E, Harris AC, Jaeger TF (2015) Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. Lang Cogn Neurosci 30:1009–1032.

[47] Zipf GK (1949) Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology (Addison-Wesley Press, Cambridge, MA).

[48] Himmelmann NP (2014) Asymmetries in the prosodic phrasing of function words: Another look at the suffixing preference. Language 90:927–960.

[49] Piantadosi ST, Tily H, Gibson E (2011) Word lengths are optimized for efficient communication. Proc Natl Acad Sci USA 108:3526–3529.

[50] Sacks H, Schegloff EA, Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. Language 50:696–735.

[51] Strunk J, Schiel F, Seifart F (2014) Untrained forced alignment of transcriptions and audio for language documentation corpora using WebMAUS. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014), eds Calzolari N, et al. (European Language Resources Association, Reykjavik, Iceland), pp 3940–3947.

[52] Kisler T, et al. (2016) BAS speech science web services–An update of current developments. Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), eds Calzolari N, et al. (European Language Resources Association, Paris, France), pp 3880–3885.

[53] Goldman-Eisler F (1968) Psycholinguistics: Experiments in Spontaneous Speech (Academic, London).

[54] Hieke AE, Kowal S, O'Connell DC (1983) The trouble with "articulatory" pauses. Lang Speech 26:203–214.

[55] Campione E, Véronis J (2002) A large-scale multilingual study of silent pause duration. Speech Prosody 2002:199–202.

[56] Lehiste I (1970) Suprasegmentals (MIT Press, Cambridge, MA).

[57] Baayen RH, Davidson DJ, Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. J Mem Lang 59:390–412.

[58] R Core Team (2018) R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna), Version 3.4.4. Available at https://www.r-project.org/. Accessed March 15, 2018.

[59] Bates D, Mächler M, Bolker BM, Walker SC (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67.

[60] Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300.

[61] Fox J (2003) Effect displays in R for generalised linear models. J Stat Softw 8:1–27.

[62] Subbiah M, Rajeswaran V (2017) Proportion: A comprehensive R package for inference on single Binomial proportion and Bayesian computations. SoftwareX 6: 36–41.