Prosody or Information rate? Informativity effects on pitch, intensity and duration in Mandarin Chinese words

Kevin Tang[1] & Jason A. Shaw[2]
[1]*University of Florida,* [2]*Yale University*

Synopsis: At least in English, informativity influences word duration, with more informative words being shorter than less informative words (Seyfarth 2014). This effect has been interpreted as phonetic detail lexicalized through a production-perception feedback loop (Wedel, 2007, 2012; Hay et al., 2015). However, the cognitive mechanism regulating the duration-predictability tradeoff in online production is still being debated (e.g., Buz & Jaeger, 2017). We evaluated two competing hypotheses in a corpus study of Mandarin Chinese, a language known to mark prosodic prominence with duration, but also pitch and intensity (e.g., Wang and Xu 2011).

Information Rate Hypothesis (IRH): Faster speech correlates with more predictable words (Cohen Priva 2016). Coupe et al. (2019) report that the average speech rate of a language varies systematically with information density, giving rise to a universal information rate (39 bits/second). Pressure to maintain a constant information rate may therefore drive the correlation between predictability and word duration. This hypothesis predicts that duration is the only phonetic parameter related to informativity. To the extent that informativity correlates with other phonetic parameters, they are expected on this hypothesis to be mediated by duration.

Smooth Signal Redundancy Hypothesis (SSRH): An alternative hypothesis is that prosodic structure functions to regulate signal redundancy (Aylett & Turk, 2004; Turk 2010). SSRH predicts that prosody is the force driving predictability effects on word duration. On this account, informativity may condition variation in duration because unpredictable words are more likely to receive prosodic prominence, e.g., as "new information". This predicts that the phonetic dimensions related to prosodic prominence may all be subjected to informativity effects.

Method: To test these hypotheses, we calculated word informativity using a 431 million word subtitle corpus and the acoustic information of content words from a speech corpus (400,000 tokens by 1,655 individuals). Linear mixed effects models were fit to measures of word duration, maximum intensity and maximum f0. The model structure is given in (1). Random effects included intercepts for item, tone sequence, and speaker, and a by-speaker random slope for informativity. The key fixed effect was (forward/backward) informativity, defined in (2), which was included along with 15 other control variables, including frequency, local predictability, and speech rate. For pitch and intensity, we also did a mediation analysis, in which we tested whether apparent effects of informativity could be explained instead by duration.

Results and Discussion: As shown in Table 1, informativity significantly influenced all three variables known to be associated with prominence in Mandarin Chinese. A unit increase in forward informativity conditioned higher maximum pitch, longer duration and higher intensity. The positive coefficients indicate that the directions of the effects are in the direction predicted by the lexicalization of prosodic prominence. Moreover, the effects of informativity on pitch and intensity remained significant (and largely unchanged) when duration was added as well to the model, suggesting that the effects of informativity on pitch and intensity were direct (c.f., Gleason & Cohen-Priva 2018). That Informativity conditioned variation in all three phonetic parameters related to prosodic prominence supports the SSRH over IRH. As far as we know, this is the first study to show that multiple phonetic dimensions associated with prosody (pitch, intensity, duration) leave a phonetic "residue" on the lexicon.

Table 1: *Summary of Informativity effects on duration, maximum pitch and maximum intensity*									
	Duration			Maximum Pitch			Maximum Intensity		
	β	**SE**	***t***	***β***	**SE**	***t***	***β***	**SE**	***t***
Forward informativity	0.0099	0.0014	7.07	0.0039	0.0008	4.66	0.4043	0.0696	5.81
Backward informativity	0.0059	0.0013	4.48	-0.0034	0.0008	-4.22	0.0813	0.0667	1.21

(1) Model structure: *Dependent variable (either Duration, Maximum intensity, Maximum pitch) ~ Frequency + Forward predictability + Backward predictability + Forward informativity + Backward informativity + Word length + Preceding disfluency + Following disfluency + Preceding pause duration + Following pause duration + Preceding speech rate + Following speech rate + Previous self-mention + Previous cross-speaker mention + Age + Gender + Syntactic category + (1 | Word type) + (1 | Tone sequence) + (1 + Forward informativity + Backward informativity | Speaker)*

(2) Informativity is defined as the negative log average contextual predictability of a word in every context in which it appears in, weighted by the contextual predictability of the contexts (Piantadosi et al. 2011). The numerical expression is given below, where *c* is a context, *C* is the set of all contexts, *w* is a word and *W* is the set of all words. For forward informativity the context was the preceding word; for backward informativity the context was the following word.

Informativity: $- \sum_{c} Pr(C = c \mid W = w) \, log_2 \, Pr(W = w \mid C = c)$

References

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; **Cohen-Priva, U.** (2017). Not so fast: Fast speech correlates with lower lexical and structural information. *Cognition*, *160*, 27-34. **Coupé, C., Oh, Y. M., Dediu, D., & Pellegrino, F.** (2019). Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances*, *5*(9), eaaw2594; **Hay, J. B., Pierrehumbert, J. B., Walker, A. J., & LaShell, P.** (2015). Tracking word frequency effects through 130years of sound change. Cognition, 139, 83-91. **Jaeger, T. F., & Buz, E.** (2017). Signal reduction and linguistic encoding. Handbook of psycholinguistics. Wiley-Blackwell. **Piantadosi, S. T., Tily, H., & Gibson, E.** (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences*, 108(9), 3526-3529; **Seyfarth, S.** (2014). Word informativity influences acoustic duration: Effects of contextual predictability on lexical representation. *Cognition*, 133(1), 140-155; **Turk, A.** (2010). Does prosodic constituency signal relative predictability? A smooth signal redundancy hypothesis. Laboratory Phonology, 1(2), 227-262; **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. **Wang, B., & Xu, Y.** (2011). Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. *Journal of Phonetics*, 39(4), 595-611. **Wedel, A. B. (2007).** Feedback and regularity in the lexicon. *Phonology*, *24*(1), 147-185. **Wedel, A. (2012).** Lexical contrast maintenance and the organization of sublexical contrast systems. Language and Cognition, 4(4), 319-355.