Functional Load modulates speech production, but not speech perception:
Evidence from Thai vowel length

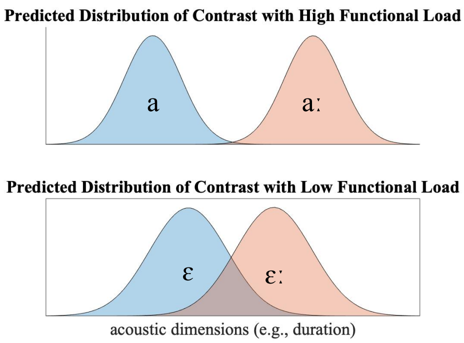
Sireemas Maspong & Francesco Burroni
Department of Linguistics, Cornell University

In this paper, we show that functional load (FL) modulates the production of short-long vowel contrasts in Bangkok Thai. However, no effect is observed in perception. FL, a measure of how much a language relies on a phonemic contrast to differentiate words ([1],[2],[3]), has been hypothesized to play a role in sound change and in the evolution of sound inventories. For instance, high FL of a phonemic contrast has been shown to correlate with a resistance to mergers ([3]). However, how do these diachronic tendencies come about in day–to–day speech? One possibility we explore in this paper is that FL has real time effects on speech production and/or perception. It has recently been reported that FL can help predict fine-grained speech production details of phonological contrasts, such as the singleton to geminate duration ratio observed in different languages ([4]). We extend this research program to vowel length in Bangkok Thai.

Research questions. The questions we investigate are the following: (i) Does FL correlate with the duration ratio of long to short vowels just like it has been shown to correlate with the ratio of long to short consonants ([4])? (ii) Does FL correlate with the amount of distributional overlap between short and long vowels of the same quality? (iii) Does FL correlate with "ease" of perceptual identification of the relevant short/long vowel contrasts, as measured by, e.g., reaction times?

Predictions. If FL has effects on speech production and perception, we predict vowel pairs, e.g., [a ~ aː], with high FL to show (i) larger duration ratios between long and short vowels, and (ii) less overlap between the distribution of the short and long vowel pair member (Fig. 1). Furthermore, if FL modulates perception, we predict that it would (iii) correlate with shorter reaction times when classifying stimuli of ambiguous duration, as listeners are more often exposed to the high FL contrast. Opposite patterns are expected to hold for contrasts with low FL.

Fig. 1 Predictions



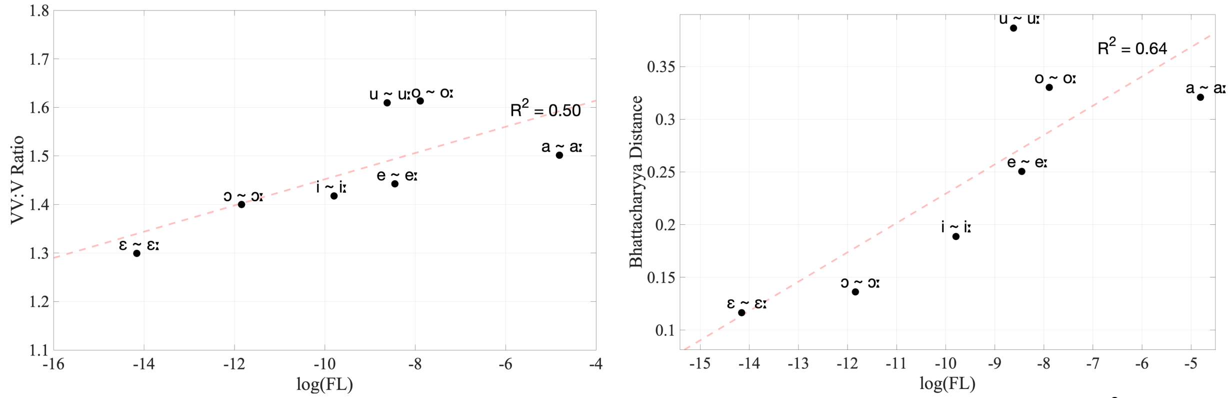
Methodology. The FL of vowel length contrast for each vowel pair was calculated using word probabilities of a unigram model based on monosyllabic forms extracted from the Thai National Corpus (~22 million tokens). The formula used to calculate FL follows [2]: $FL(x,y) = \frac{H(L)-H(L_{xy})}{H(L)}$, where $FL(x,y)$ is the FL of the contrast (x,y), $H(L)$ is the baseline entropy of the Thai unigram model, and $H(L_{xy})$ is the entropy of a Thai unigram model where the target contrast (x,y) has been neutralized.

Production. Thai was chosen as a case study because it shows vowel length contrast for all monophthongs and duration, rather than vowel quality, is the primary cue for this contrast ([5]). Data were collected from 20 native Bangkok Thai speakers (12F;8M). The stimuli were 189 attested Thai words with all licit phonotactic and tone combinations. Target words were embedded in a frame sentence and presented in Thai orthography. Participants were asked to produce each sentence three times. Carrier sentences were force-aligned, inspected by a native Thai speaker, and corrected when necessary. Vowel duration of target words was extracted. The grand mean of the duration of each vowel was used to calculate the duration ratio of the short/long vowel pair of the same quality. Furthermore, the durational distributions of each vowel pair were analyzed to obtain their Bhattacharyya Distance, used to operationalize the distance (or less overlap) between the two probability distributions ([6]). We calculated the correlation of (i) the duration ratio and *log*(FL) and (ii) the Bhattacharyya Distance and *log*(FL) using Pearson's r and Kendall τ rank correlations, the choice of using *log*(FL) and of the statistical tests follows [4].

Perception. 15 native Thai speakers completed an ABX discrimination task. The stimuli were nonce words of shape [f (i/u/ɛ/ɔ/a) p] (chosen to yield nonce words). We modified vowel duration of the stimuli into 11 steps. To prevent possible spectral biases, we used two types of stimuli: lengthened short vowel and shortened long vowel. The modification was 10 ms at each step. Every unique stimulus was presented 3 times. The grand means of the reaction time for step 4 to 8, the region of the continuum where the responses as short vowels for all vowels are in the range of 10-90 percent, Fig. 3 (left), were extracted and analyzed for their correlation with *log*(FL) using Pearson's r and Kendall τ rank correlations.

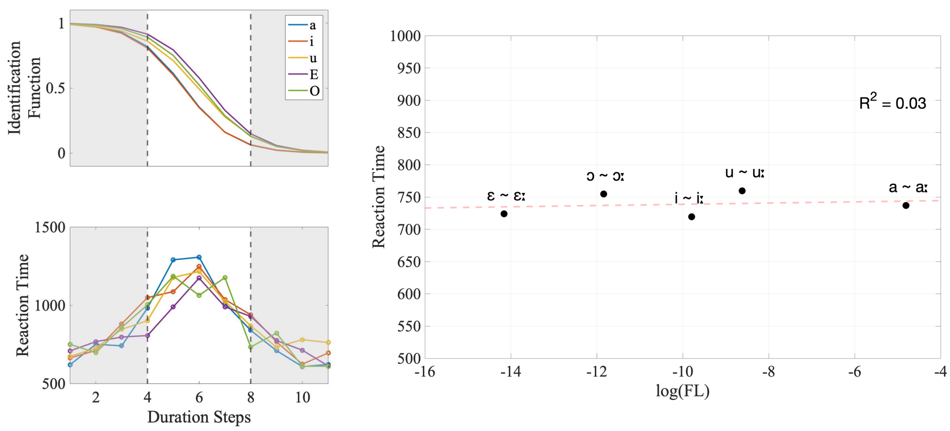
Results. *Production*. Duration ratio displays a correlation with *log*(FL) (ρ = .7; τ = .71; R^2 = .5), Fig 2. (left), Bhattacharyya Distance also displays a correlation with *log*(FL) (ρ = .8; τ = .61; R^2 = .64), Fig 2. (right).

Fig. 2 Correlations of FL and duration ratio (left)
and correlations of FL and Bhattacharyya Distance (right)



Perception. Mean reaction times *do not* show correlation with *log*(FL) (ρ = .18; τ = .2; R^2 = .03), Fig. 3 (right).

Fig 3. Mean reaction time for each step (left)
and correlations of FL and mean reaction time of stimuli with ambiguous duration (right)



Conclusion. In sum, the production results for both duration ratio and Bhattacharyya Distance conform with the predictions that FL modulates the production of short and long vowel contrasts. However, this is not the case for perception where, for instance, [a~a]: – the contrast with highest FL – has comparable mean reaction time with contrast that have much lower FL, such as [ɛ~ɛː]. We evaluate the consequences of our findings for theories of sound change that are often based on perception alone [7] and discuss why FL does not affect perception in the same way it affects production. Finally, we sketch out how FL may be accommodated in a dynamical model of speech production, like the task dynamic model of Articulatory Phonology, by modulating gestural durations on the basis of feedback [8].

References

[1] Hockett, Charles F. 1967. The Quantification of Functional Load. *Word*. Routledge 23(1–3). 300–320.

[2] Surendran, Dinoj & Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In Ole Nedergaard Thomsen (ed.), *Competing Models of Linguistics Change: Evolution and Beyond. In commemoration of Eugenio Coseriu (1921-2002)*, 43–58. Amsterdam & Philadelphia: Benjamins.

[3] Wedel, Andrew, Abby Kaplan & Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2). 179–186.

[4] Tang, Kevin & John Harris. 2014. A functional-load account of geminate contrastiveness: a meta-study. Presented at the Linguistics Association of Great Britain 2014.

[5] Abramson, Arthur S. & Nianqi Reo. 1990. Distinctive vowel length: Duration vs. spectrum in Thai. *Journal of Phonetics* 18(2). 79–92.

[6] Bhattacharyya, A. 1943. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society* 35. 99–109.

[7] Ohala, John J., The phonetics of sound change. In Charles Jones (ed.), *Historical Linguistics: Problems and Perspectives*. London: Longman. 237-278.

[8] Tilsen. Sam. 2016. Selection and coordination: The articulatory basis for the emergence of phonological structure. *Journal of Phonetics* 55, 53-77.