Incorporating continuity in phonological models of the syllable using NAP

Aviad Albert[1] & Bruno Nicenboim[2]
[1]*University of Cologne*, [2]*Tilburg University*

Prevailing sonority models are often based on the *Sonority Sequencing Principle* (SSP; [1]) employing discrete segmental entities and slopes that result from their concatenation in symbolic time to determine syllabic well-formedness. While evidently useful, the type of discrete computation underlined by such models is incompatible with many cognitive models of the mind [2], while also exhibiting gaps in empirical coverage (e.g. /s/-stop clusters [3]). We present a new model of sonority that incorporates continuous entities and dynamic procedures to account for syllabic well-formedness. The *Nucleus Attraction Principle* (NAP) models a bottom-up route of linguistic inferences, mapping information from acoustics to perception and cognition in a functionally-motivated manner. Importantly, NAP requires fewer formal assumptions compared to SSP-based models, while at the same time exhibiting a better empirical coverage. We present a perception experiment that was designed to test NAP against four variants of SSP-based models. We use a Bayesian data analysis approach to test and compare the different sonority models [4], where NAP is found to be superior.

NAP embraces the most basic postulate of sonority-based models, whereby the most sonorous element is contained within the nucleus of the syllable. However, instead of adding further formal assumptions about sonority slopes in symbolic time, we simply model the link between sonority and syllabic nuclei as a dynamic process in real time, whereby all the portions of the speech signal compete against each other for the nucleus. Thus, sonority is the quality that *attracts* the nucleus such that the most sonorous portion in a given sequence is predicted to win the competition for the nucleus. Syllabic ill-formedness is therefore directly related to the degree of nucleus competition that a given syllabified portion incurs.

We quantify the competition potential of continuous speech portions by obtaining *periodic energy* from the acoustic signal, as a measure of the acoustic intensity of the pitch-bearing component of the signal, which we consider to be the most appropriate correlate of sonority [5,6]. We calculate the *Center of Mass* of the area under the periodic energy curve at two locations to measure the displacement of energy to the left of the syllable in order to estimate the competition potential between the (losing) onset and the (winning) nucleus (Figure 1).

We recorded 29 complex onset clusters in a /CCal/ word frame, alongside /CəCal/ and /əCCal/ fillers (produced in a non-final position of a broad-focus sentence), and we asked 51 native German speakers to determine if they hear one or two syllables, echoing Berent et al.'s experimental paradigm (e.g. [7]). We logged the reaction time (RT) of all "1 syllable" responses to single vowel targets as a measure of the processing cost associated with syllabifying these stimuli with a single nucleus. Sonority models were tested on their ability to predict these RTs given the cluster. We consider 4 SSP-based models, using 2 different sonority hierarchies—featuring a collapsed class of obstruents (*col*) and an expanded class of obstruents (*exp*) that includes distinctions between voiced vs. voiceless, and stops vs. fricatives—with 2 types of sonority principles: the SSP and the *Minimum Sonority Distance* (MSD; [8]), which favours steeper rising sonority slopes over shallower rises.

In the statistical models [9-11], we accounted for the ordinal form of SSP-based model predictions vs. the continuous scores in NAP, and we fitted a *null* model as baseline. All sonority models exhibited a clear effect on measured RTs, suggesting that they all capture at least some important aspects of sonority (e.g. Figures 2a-b). For model comparison, we used k-fold (k=15) cross validation stratified by subjects, where NAP was found to have superior predictive accuracy indicating a stronger ability to generalize to unseen stimuli (Table 1).

To conclude, the NAP model presents methodological, theoretical and empirical advantages over prevailing SSP-based models, supporting the incorporation of continuity in phonology and validating the choice of periodic energy as the acoustic correlate of sonority.

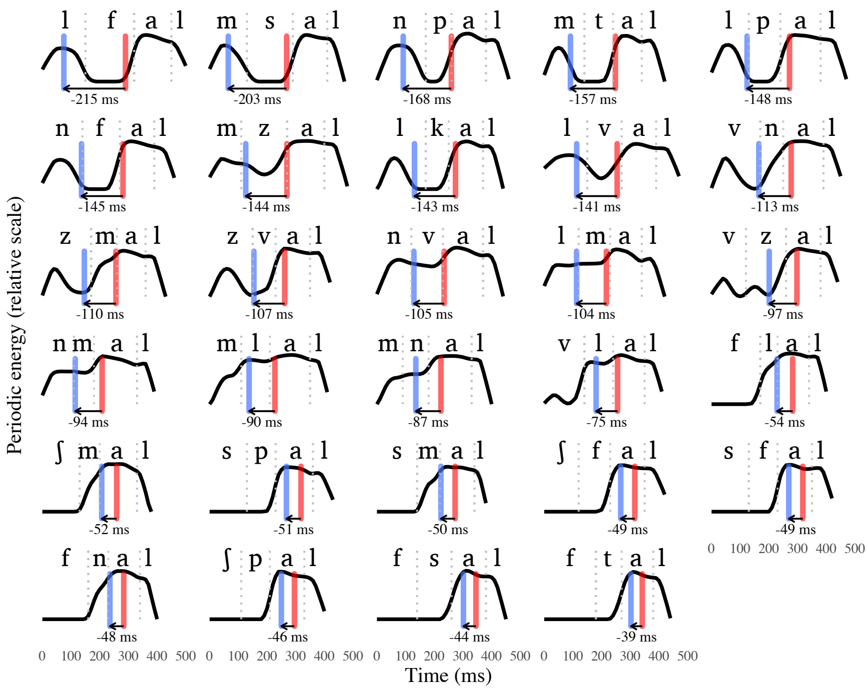
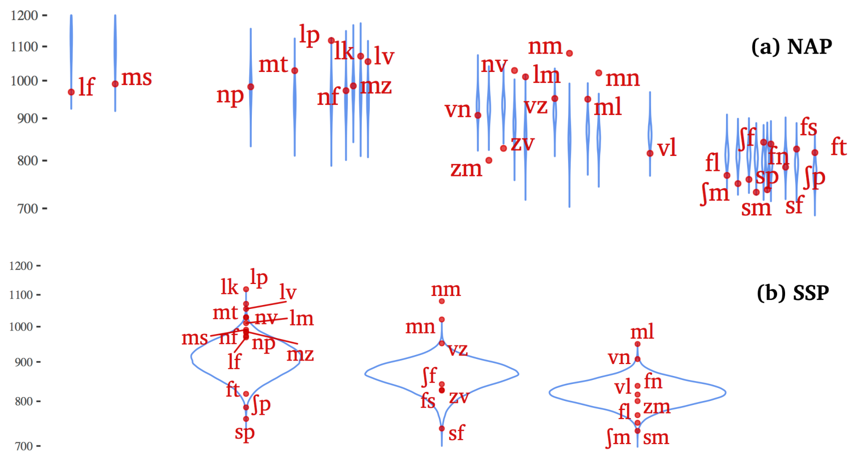


Figure 1. Smoothed periodic energy curves (black) of the 29 experimental targets. Red vertical lines denote the center of mass of the entire syllable, blue vertical lines denote the center of mass of the left portion (from the beginning up to the red line). Distances between mass centers (black arrows and corresponding numbers) denote competition potentials, i.e. NAP's well-formedness scores. Grey dotted vertical lines denote manual segmentation (for exposition purposes only).



Figures 2(a-b). **Model fit examples:** Observed mean log-transformed response times (y-axis) are depicted with red points; distribution of simulated means based on individual models are depicted with blue violins. Order of stimuli on the x-axis (from less to more well-formed) reflects model scores: scalar for NAP (2a) and ordinal for SSP (2b).

Model	elpd	Difference in elpd	Difference SE	Weight
NAP	-28680	0.00	0.00	0.77
MSD$_{\text{exp}}$	-28800	-120.00	18.83	0.1
SSP$_{\text{exp}}$	-28818	-138.78	19.68	≈ 0
MSD$_{\text{col}}$	-28833	-153.72	21.75	0.06
SSP$_{\text{col}}$	-28838	-157.79	21.95	0.07
Null	-28853	-173.71	19.92	≈ 0

Table 1. **Model comparison:** The table is ordered by the expected log-predictive density (elpd) score of the models (higher score indicating better predictive accuracy). SE = Standard Error. Weights represent the combined weights of the individual models that maximize the total elpd score.

References: **[1]** Parker, S. 2017. Sounding out sonority. *Language and Linguistics Compass* 11(9); **[2]** Spivey, M. 2007. *The Continuity of Mind*. New York: Oxford University Press; **[3]** Goad, H. 2016. Sonority and the unusual behaviour of /s/. In *Challenging Sonority: Cross-Linguistic Evidence*. Ed. M.J. Ball & N. Müller. Sheffield; Bristol: Equinox Publishing; **[4]** Nicenboim, B. & Vasishth, S. 2016. Statistical methods for linguistic research: Foundational Ideas - Part II. *Language and Linguistics Compass* 10(11), 591–613; **[5]** Deshmukh, O., J. Singh & C. Espy-Wilson. 2004. A novel method for computation of periodicity, aperiodicity and pitch of speech signals. *IEEE Xplore* 1:I.117-I.120; **[6]** Lass, R. 1984. *Phonology: An Introduction to Basic Concepts*. Cambridge Textbooks in Linguistics. Cambridge: University Press; **[7]** Berent, I., T. Lennertz, J. Jun, M. A. Moreno, and P. Smolensky. 2008. Language universals in human brains. *PNAS* 105 (14): 5321-5325; **[8]** Steriade, D. 1982. *Greek prosodies and the nature of syllabification*. PhD dissertation, Massachusetts Institute of Technology. **[9]** R Core Team. 2018. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing; **[10]** Stan Development Team. 2018. *Stan: A C++ library for probability and sampling*; **[11]** Bürkner, P. C. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1), 1–28.