Cross-linguistic deviations from the sonority sequencing are perceptual-articulatory effects

Background: The sonority sequencing principle (SSP) is a fundamental governing principle of the syllable structure (Clements, 1990). Under Optimality Theory, sonority constraints are imposed universally on languages, as formalised in *SonSeq (Kager, 1999) or in fixed sonority constraint hierarchy realised through local conjunction in a local domain like $[\text{*Ons/gl+*Ons/obs}]_{ons}$ >>... >>$[\text{*Ons/obs+*Ons/gl}]_{ons}$ (Prince & Smolensky, 1993/2004). These sonority constraints predict if a language allows marked SSP-violating clusters, it also allows unmarked SSP-conforming ones. Cross-linguistically, the attestation of marked SSP-violating clusters is the result of intervening constraints that dominate sonority constraints (Zec, 2007) or due to extra-syllabic/-prosodic status of the segment responsible for the violation (e.g., /s/ in sC) (Clements & Keyser, 1983). However, the sonority role for syllabification continues to be tested by its lack of definition (Parker, 2002) and by clusters that violate the sonority sequencing, especially clusters that cross over sonorants and obstruents. An increasing number of clusters that run counter to the SSP are reported from under-studied languages recently, and the establishment of complex syllable structures where all sonority profiles break down to form an important typological syllable category has put a serious challenge to the sonority role as a governing principle of syllabification (Easterday, 2019). It has remained unclear how the current sonority theory couched in different phonological theories can capture the diverse sonority profiles found cross-linguistically. To this aim, the current study aims to clarify the empirical status of sonority sequencing and examine the cross-linguistic standing of sonority constraints in large typological corpora.

Methods: 1) Permissible consonant clusters in each of the cross-linguistically diverse 496 languages were obtained from two large lexical databases, CLICS^2 (List et al., 2018) and AusPhon (Round, 2017); 2) all long consonant clusters larger than two (CCC, CCCCC, etc.) were broken up into diclusters, e.g., /mc'vrt-/ into /mc'-, c'v-, vr-, rt-/; 3) permissible sonority contours in each language were calculated by adopting a scale ([gl>nas>liq>obs]) in which sonority is defined through binary features (Clements, 1990); 4) SSP-conforming, sonority plateau, and SSP-violating language and consonant cluster attestations were counted, respectively, and then the associations between two adjacent sonority classes were examined.

Results: Of a total of 496 languages and 4189 consonant clusters studied, SSP-conforming languages and clusters are attested in more languages and clusters, followed by sonority plateaus, and then by -violations. SSP-conformers, -plateaus, and -violators account for 51.0%, 30.3%, and 18.7% of 496 languages, and 46.2%, 42.8%, and 11.1% of 4189 clusters, respectively. There is also a significant association between two adjacent clusters, counted either in cluster attestations ($\chi^2(9, N=4189)=271.71, p<.001$) or in language attestations ($\chi^2(9, N=875)=68.67, p<.001$) (see Figure 1 below), suggesting given some cluster, some particular cluster is more (or less) likely to occur. Specifically, perceptually or articulatorily favoured clusters are more likely to occur, whereas disfavoured ones are less likely so.

Discussion: SSP violations are found to be common, however, the SSP is still found to be a cross-linguistic tendency. This cross-linguistic tendency of sonority sequencing is well in line with the prediction of sonority constraints couched in OT, as captured in Figure 2 below. Meanwhile, the study also shows some marked clusters like OL (compared with OG) or LG (compared with all SSP-conforming or -plateau clusters) are statistically more likely (see Figure 2 and Figure 3 below), while some unmarked clusters like GG, LL, OO or NL (compared with all SSP-violating clusters) are statistically less likely to occur. Among these, some have shown that they are articulatorily (LG) (Jun, 1995) or perceptually (OL) favoured (Ohala & Kawasaki-Fukumori, 1997), while some like NL (Henke et al., 2012), or GG, LL or OO (OCP-type constraints (Goldsmith, 1990)) are perceptually or articulatorily disfavoured. Therefore, while the SSP stands as a tendency, the current study also indicates perceptual and articulatory roles in shaping cross-linguistic phonotactic deviations from sonority sequencing.

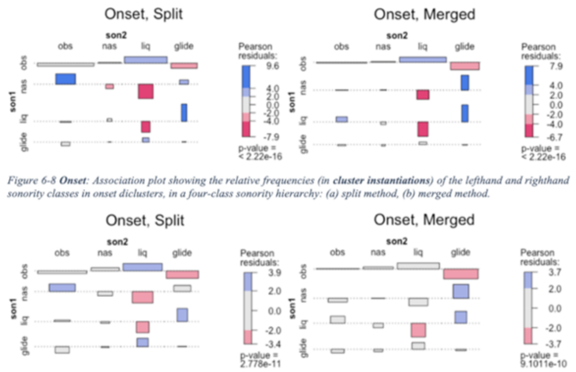

Figure 1 Association plot showing the relative frequencies of consonant cluster attestations on the upper row, and language attestations on the lower row by adopting a four-level sonority hierarchy (gl > liq > nas > obs). The plots on the left show the results when affricates and prenasalised stops are treated as sequences /t s, m b/ (split), and those on the right show the results when these two types of complex segments are treated as one segmental unit/t͡s, ^{m}b/ (merged).

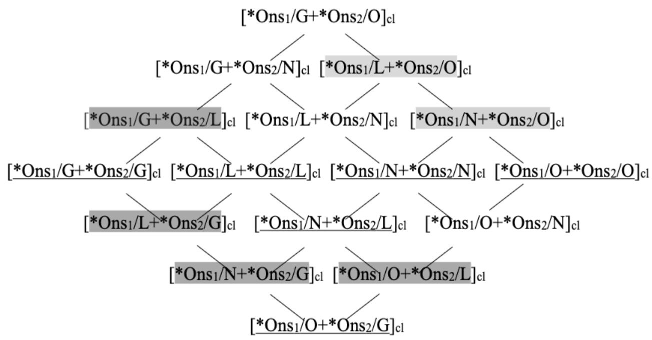

Figure 2 Sonority constraints on consonant clusters showing the relative (un)markedness status of clusters, derived from adopting a four-level sonority hierarchy, glides > liquids > nasals > obstruents, realised through harmonic alignment of two sonority scales in a local domain where the domain is defined as cluster following (Smolensky, 2006). Greyed clusters are clusters that are statistically more likely to occur, cross-linguistically; dark grey stands for clusters that are more likely to occur regardless of whether affricates and prenasalised are assumed to form one segment or not, and light grey stands for the clusters that are likely to occur only when these two types of complex segments are assumed to form one segment. Underlined clusters are the ones that are statistically less likely to occur, cross-linguistically. (Acronyms are coded as O: obstruents, N: nasals, L: liquids, G: glides).

References

Clements, G. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and physics of speech* (pp. 283-333). Cambridge: Cambridge University Press.

Clements, G., & Keyser, S. (1983). *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.

Easterday, S. (2019). *Highly complex syllable structure: A typological and diachronic study*. Berlin: Language Science Press.

Goldsmith, J. (1990). *Autosegmental and metrical phonology* (Vol. 1). Basil Blackwell.

Henke, E., Kaisse, E., & Wright, R. (2012). Is the sonority sequencing principle an epiphenomenon? In S. Parker (Ed.), *The Sonority Controversy*. Berlin & Boston: De Gruyter Mouton.

Jun, J. (1995). *Perceptual and articulatory factors in place assimilation: An Optimality Theoretic approach.* (Ph.D. dissertation). University of California, Los Angeles, CA.

Kager, R. (1999). *Optimality Theory*. Cambridge Cambridge University Press.

List, J.-M., Greenhill, S., Anderson, C., Mayer, T., Tresoldi, T., & Forkel, R. (2018). Database of cross-linguistic colexifications.

Ohala, J., & Kawasaki-Fukumori, H. (1997). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In S. Eliasson & E. Jahr (Eds.), *Language and its Ecology* (Vol. 2, pp. 343-366). Berlin & New York: De Gruyter Mouton.

Parker, S. (2002). *Quantifying the sonority hierarchy.* (Ph.D. dissertation). University of Massachusetts at Amherst, Amherst, MA.

Prince, A., & Smolensky, P. (1993/2004). *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA, & Oxford: Blackwell.

Round, E. (2017). *The AusPhon-Lexicon project: 2 million normalised segments across 300 Australian languages*. Paper presented at the 47th Poznan Linguistic Meeting, Poznan.

Zec, D. (2007). The syllable. In P. de Lacy (Ed.), *Handbook of Phonology* (pp. 161-194). Cambridge: Cambridge University Press.