The Hitchhiker's Guide to Harmony Interactions

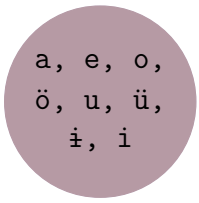
Alëna Aksënova

Keywords: harmonies, typology, mathematical linguistics, learnability

Overview In this paper, I am considering harmony systems with multiple feature spreadings as a litmus test for the possible configurations of items involved in certain long-distance dependencies within the same language. The question that is being asked is what are the relations between the sets of items that take part in these spreadings. The logically possible configurations are the following ones: *same set*, when both spreadings operate over the same elements, *distinct sets*, when the sets of items involved in the harmonies are completely different, *set and its subset*, when one harmony operates over the subset of items involved in another harmony, and *incomparable sets*, when the sets of items involved in the two harmonies only partially overlap. I start by investigating typology, and show that the fourth option – the overlapping sets – is the only one that appears to be unattested. While this could be surprising, further I discuss formal reasons that explain such restriction. Going forward, considering such overlapping sets would produce the blow-up of options of the possible harmonic sets configurations.

Typology In this subsecton, I present typological data showing each of the configurations listed above, except the last one – the one that, to the best of my knowledge, is unattested. Here and further, by *items involved in a harmony* I mean items that are not transparent for the harmonic process, i.e. they are either blockers, or undergoers.

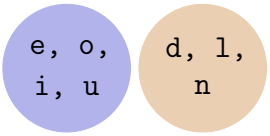
Same set Many harmonies with several feature spreadings operate over the same set of elements. This does not mean that undergoers and blockers are the same for both harmonies, it only means that none of the items taking part in one harmony is irrelevant for the other one. As an example of such language consider Yakut (TURKIC), (Sasa 2009). In this language, the harmonizing features are [front] and [round]. All vowels must agree in fronting (1-4). However, labial harmony spreads from low vowels onto both low and high ones (1,2), from high vowels to high ones (3), but it cannot spread from high vowels to low ones (4).



Yakut

(1) **oɣo**-l**o**r 'child-PL' (3) m**u**r**u**m-**u** 'nose-ACC'
(2) **oɣo**-n**u** 'child-ACC' (4) t**ü**nn**ü**k-l**e**r 'window-PL'

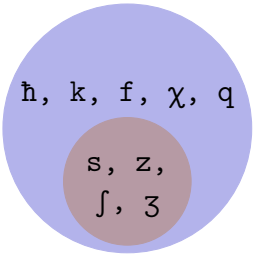
Distinct sets In some cases, harmonies operate over completely different sets of elements. For example, in Kikongo (BANTU), vowel harmony co-exists with consonant assimilation, (Ao 1991, Hyman 1998). Vowels must agree in their [height] specification within a word (5-8). Suffixal consonants that are /d/ or /l/ underlyingly (5,6) are realized as /n/ if they are preceded by /n/ in the stem (7,8).



Kikongo

(5) -suk-i**d**i- 'wash-PERF.ACT' (7) -me**n**g-e**n**e- 'hate-PERF.ACT'
(6) -suk-u**l**u- 'wash-PERF.PASS' (8) -me**n**g-o**n**o- 'hate-PERF.PASS'

Set and its subset The only remaining possibility among the attested ones is the case when one harmony operates over the subset of items involved in another harmony. In Imdlawn Tashlhiyt (BERBER), prefixal sibilants harmonize with the stem ones in voicing and anteriority, (Hansson 2010, McMullin 2016). Whereas nothing blocks the anteriority harmony (9-12), the voicing assimilation is blocked by any intervening voiceless obstruents (11,12).



Imdlawn Tashlhiyt

(9) **s**-a**s:**twa 'CAUS-settle' (11) **s:**-u**kz** 'CAUS-recognize'
(10) **ʒ**-m:**ʒ**dawl 'CAUS-stumble' (12) **ʃ**-**q**u**ʒ:**i 'CAUS-be.dislocated'

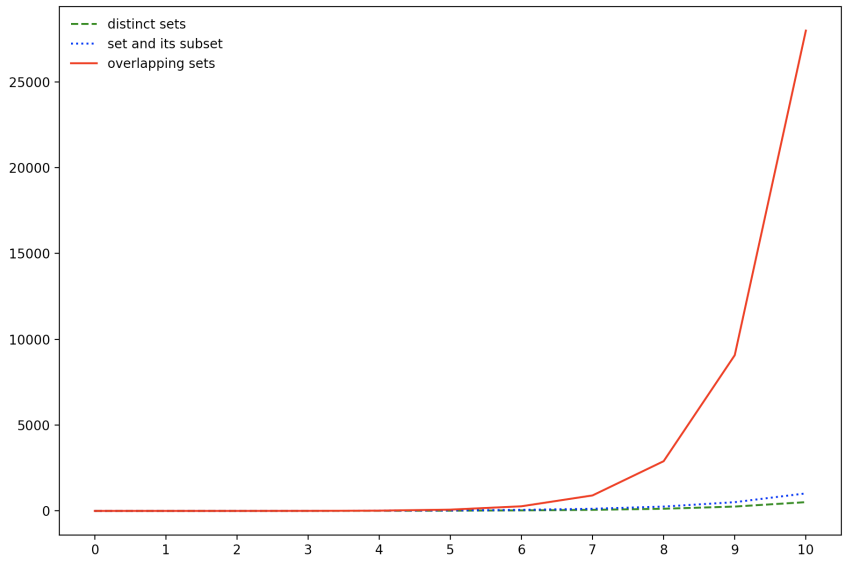
To the best of my knowledge, there are no languages that have multiple spreadings that operate

over only partially overlapping sets of elements.

Partitioning Here, I discuss the problem of configurations of harmonic sets from a formal point of view. Namely, I will show that if we consider all possible partitioning of a set into two subsets, then the vast majority of the resulting sets are the partially overlapping ones, and it is exactly the configuration that seems to be absent from natural languages. Eliminating the option of partially overlapping sets helps to simplify the variety of choices for natural language harmonic systems configurations. Including this option causes a blow-up of the possible harmonic systems: see below for the concrete numbers and formulas. This result hightlights the importance of restricting the system in a way that natural languages restrict themselves.

`Set and its subset` For a set of n elements, the amount of its subset is equal to choosing k elements from a set of n, or $\binom{n}{k}$. Two subsets need to be excluded: $k = 0$, where one of the sets is empty, and $k = n$, where the two sets are equivalent (the 'same set' case). The amount of all other proper subsets is given by $\sum_{k=1}^{n-1} \binom{n}{k} = 2^n - 2$.

`Distinct sets` The general case of partitioning a set of n elements into k disjoint subsets is given by Stirling Numbers of the Second Kind or $S(n,k)$, see (Knuth 1968). For the partitioning into 2 sets, the following formula can be used: $S(n,2) = \frac{1}{2}\sum_{j=0}^{2}(-1)^{2-j}\binom{2}{j}j^n$.



`Overlapping sets` In this case we want to partition a set of n elements into two sets with a non-empty intersection. This problem can be divided into two subproblems: partitioning the set of n elements into 3 disjoint sets using $S(n,3)$; and ordering the partitions to generate all possible intersections. The solution can be obtained by the following formula: $3 * S(n,3) = \frac{1}{2}\sum_{j=0}^{3}(-1)^{3-j}\binom{3}{j}j^n$.

The picture above illustrates growth of these functions. For example, for 10 elements, there are 511 ways to arrange them in two distinct sets, 1022 ways to obtain a set and its subset, and 27990 ways to create two overlapping sets.

Discussion I showed that there are no attested cases when multiple harmonies within the same language operate over only partially overlapping sets of elements – moreover, this typologically unattested case would lead to a blow-up of possible configurations of harmonic systems.
This is just preliminary research about the typology of long-distance processes and the math behind it, and, of course, a lot is still remained unexplored. However, this result can be interesting from several different perspectives. First, it reveals new typological generalizations about harmonic systems and natural languages in general. Secondly, it might shed light on the issues related to the learnability of long-distance processes. And, lastly, it brings a desired naturalness to the models of natural languages.

Ao, B. 1991. Kikongo nasal harmony and context-sensitive underspecification. *LI* 22(1):193-196. ● **Hansson**, G. 2010. Long-distance voicing assimilation in Berber: spreading and/or agreement? In *Proceedings of the 2010 CLA meeting*. ● **Hyuman**, L. 1998. Positional prominence and the 'prosodic trough' in Yaka. *Phonology* 15:14-75. ● **Knuth**, D. 1968. *Fundamental Algorithms* Addison-Wesley, Reading, MA. ● **McMullin**, K. 2016. *Tier-based locality in long-distance phonotactics: learnability and typology* PhD thesis, UBC. ● **Sasa**, T. 2009. *Treatment of vowel harmony in optimality theory* PhD thesis, UIowa.