Computing Vowel Harmony:
The Generative Capacity of *Search and Copy*

Search and Copy (SC, Nevins, 2010; Mailhot and Reiss, 2007; Samuels, 2009a,b) is a procedural model of vowel harmony (VH) in which underspecified vowels trigger *searches* for *targets* that provide them with features. We analyze the weak generative capacity of SC. When search proceeds in a single direction, SC can model a phonological process only if it can be represented as a *tier-based input strictly local* function (TISL, Chandlee, 2014; Hao and Andersson, pear). The composition of multiple VH rules is however not TISL. When search proceeds in both directions, SC can model non-finite-state functions.

Background: VH is traditionally thought of as a process where features spread from a fully-specified vowel (the *donor*) to underspecified vowels (*recipients*). For example, Turkish has suffixes with vowels underspecified for backness. These vowels agree in backness with the root, e.g. the vowel /A/ in the suffixes /-lAr/ and /-A/, where A is the feature bundle [∘BACK, −HIGH, −ROUND].

(1)

	'house'	'horse'	'address'
SG	/ev/ ⇝ [ev]	/at/ ⇝ [at]	/adres/ ⇝ [adres]
PL	/ev-lAr/ ⇝ [ev-ler]	/at-lAr/ ⇝ [atlar]	/adres-lAr/ ⇝ [adres-ler]
PL.LOC	/ev-lAr-A/ ⇝ [ev-ler-e]	/at-lAr-A/ ⇝ [at-lar-a]	/adres-lAr-A/ ⇝ [adres-ler-e]

In rule-based phonology, harmony can be implemented with an *iterative output-based rule* like (2a) or a *simultaneous input-based rule* like (2b). Both formulations are *donor-based*: the backness feature is spread from the donor to the recipient.

(2) a. $A \rightarrow [\alpha\text{BACK}] \ / \ [\alpha\text{BACK}]C^* \ \_$ b. $A \rightarrow [\alpha\text{BACK}] \ / \ [\alpha\text{BACK}]\{C, A\}^* \ \_$

SC replaces the notion of iterative and simultaneous rules with a *search* operation. For an underlying form like /adres-lA_1r-A_2/, SC computes the surface form [adres-ler-e] as follows. First, each of the two As initiates a search for a vowel with the feature [αBACK]. A *direction parameter* specific to this process indicates that the search should proceed leftward. The search terminates when the first vowel with the desired features is found, and the value [αBACK] carried by this vowel is *copied* to the A initiating the search.

Tier-Based Input Strictly Local Functions: Let Σ be an alphabet, and let $\Gamma \subseteq \Sigma$ be a *tier*. A function $f : \Sigma^* \rightarrow \Sigma^*$ is *k-input strictly local on tier Γ* (k-TISL), where $k > 0$, if f is computed by a sub-sequential finite-state transducer (SFST) whose states encode the $k-1$ most recent symbols in the input stream, ignoring symbols not in Γ. To see that SC models TISL functions, let us consider the computation of /adres-lA_1r-A_2/ ⇝ [adres-ler-e] 'at the addresses' in (1). According to SC, each of the two As searches for an appropriate donor vowel, and both searches end at the /e/. In fact, since this /e/ is the last vowel with an [αBACK] feature specification, every search triggered by an A terminates at this e. Therefore, SC an be simulated using an SFST T as follows. Each state represents the identity of the most recent symbol with an [αBACK] specification. Every A encountered by T is assigned the value of [αBACK] indicated by its current state, and all other symbols are copied to the output stream faithfully. The function computed by T is then 2-TISL on tier Γ, where Γ is the set of all [±BACK] vowels. This corresponds to the set of fully specified vowels because only fully specified vowels have [±BACK] in Turkish.

Bidirectional *Search and Copy* In Woleaian, the verbal theme vowel /-A-/ is underspecified for [±LOW]. It surfaces as [e] if the closest vowels on *both* sides are [−LOW], and it surfaces as [a] otherwise. (3) illustrates this using data from Howard (1972).

(3)

1SG	/ülüm-A-ji/ ⇝ [ülümej]	3SG	/ülüm-A-la/ ⇝ [ülümal]

Nevins (2010) analyzes Woleaian as a bidirectional search initiated by the A. When the A is equidistant from two appropriate donors, the copying step is successful only if both donors contribute the same feature value to the recipient. In the 3SG from in (3), the copying step is said to fail, and the [±HIGH] receives a default valuation of [−HIGH]. In general, this predicts that in a verb of the form /V_1C^*-A-C^*V_2/, the A harmonizes with both V_1 or V_2. Such a system is not subsequential, but is Weakly Determinstic (Gainor et al., 2012; Heinz and Lai, 2013) This is represented as a leftward search that copies the [+HIGH] feature if it is found, followed by a rightward search that copies the [−HIGH] feature if it is found.

Multiple harmonies: We have shown that individual spreading processes modeled by unidirectional SC are TISL. However, since the TISL functions are not closed under composition (Chandlee et al., 2018), two harmony processes applied sequentially may result in a non-TISL mapping. For example, in addition to the backness harmony process shown in (1), Turkish high vowels harmonize in roundness with previous vowels. The genitive suffix is /-In/, where /I/ is an underspecified vowel bearing the features [∘BACK, −HIGH, ∘ROUND]. In the singular, /-In/ gets its backness and roundness features from the root: *sap-in, son-un*; but in the plural, it gets its backness feature from the plural suffix: *sap-lar-in son-lar-in, *son-lar-un*. Thus, any monolithic TISL implementation of the two harmony systems would require A to be on the tier, which contradicts the fact that backness harmony requires A *not* to be on the tier.

(4)

	'stalk'			'end'		
SG	/sap/	⇝	[sap]	/son/	⇝	[son]
PL	/sap-lAr/	⇝	[sap-lar]	/son-lAr/	⇝	[son-lar]
GEN	/sap-In/	⇝	[sap-in]	/son-In/	⇝	[son-un]
PL.GEN	/sap-lAr-In/	⇝	[sap-lar-in]	/son-lAr-In/	⇝	[son-lar-in]

Conclusion: We have shown that unidirectional SC is TISL and that bidirectional SC is not finite-state. This result improves Gainor et al.'s (2012) sub-sequential bound on unidirectional harmony systems, since Nevins (2010) analyzes all such systems using unidirectional SC. Because the composition of multiple VH patterns is not TISL, this suggests the use of *multiple* TISL functions or *multiple* tiers (Aksënova 2018).

References: **[1]** Chandlee, J. (2014). Strictly Local Phonological Processes. PhD Dissertation, University of Delaware, Newark, DE. **[2]** Chandlee, J., J. Heinz, and A. Jardine (2018, June). Input Strictly Local opaque maps. Phonology 35(2), 171–205. **[3]** Gainor, B., R. Lai, and J. Heinz (2012). Computational Characterizations of Vowel Harmony Patterns and Pathologies. In Proceedings of the 29th West Coast Conference on Formal Linguistics, Somerville, MA, pp. 63–71. Cascadilla Proceedings Project. **[4]** Hao, Y. and S. Andersson (To appear). Unbounded Stress in Subregular Phonology. In Proceedings of the 16th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology, Florence, Italy. Association for Computational Linguistics. **[5]** Heinz, J. and R. Lai (2013, August). Vowel Harmony and Subsequentiality. In Proceedings of the 13th Meeting on the Mathematics of Language (MoL 13), Sofia, Bulgaria, pp. 52–63. Association for Computational Linguistics. **[6]** Howard, I. (1972, September). A Directional Theory of Rule Application in Phonology. PhD Dissertation, Massachusetts Institute of Technology, Cambridge, MA. **[7]** Mailhot, F. and C. Reiss (2007, November). Computing Long-Distance Dependencies in Vowel Harmony. BIOLINGUISTICS 1, 028–048. **[8]** Nevins, A. (2010). Locality in vowel harmony, Volume 55. Mit Press. **[9]** Samuels, B. D. (2009a). Structure & specification in harmony. In Proceedings of the Thity-Eighth Annual Meeting of the North East Linguistics Society, Volume 2, Amherst, MA, pp. 283–296. GLSA Publications. **[10]** Samuels, B. D. (2009b, April). The Structure of Phonological Theory. PhD Dissertation, Harvard University, Cambridge, MA.