Extending TSL to Account for Interactions of Local and Non-Local Constraints

Extending TSL to Account for Interactions of Local and Non-Local Constraints

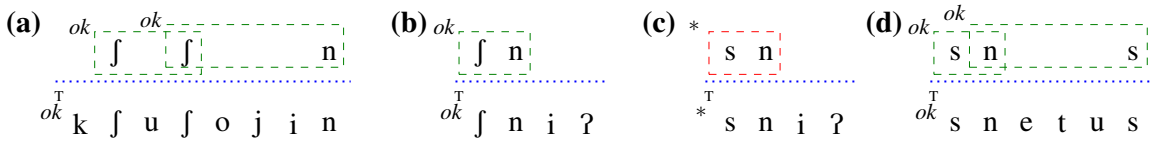
Aniello De Santo
Department of Linguistics, Stony Brook University

Keywords: phonology, phonotactics, locality, subregular hypothesis, formal language theory

Overview Recent research in computational linguistics suggests that unbounded dependencies in phonotactics, morphology, and even syntax can all be captured by the class of Tier-based Strictly Local languages (TSL). Here, I explore a new class of subregular languages obtained by relaxing a particular constraint on the tier-projection mechanism of TSL grammars. I show how a small change in the definition of tier allows us to cover phonotactic patterns that escape the standard tier-based account, without losing any of the essential properties of TSL.

TSL Phonology Many dependencies in phonology can be captured by *local constraints* that only make distinctions on the basis of contiguous substrings of segments up to some length k. For example, a (k=2) local dependency requiring /s/ to surface as [z] when immediately followed by [l] can be captured by a grammar that only contains the sequence *zl* (or, alternatively, the negative constraint *sl). Dependencies that are not locally bounded do not fit this pattern. Heinz *et al.* (2011) argue that long-distance dependencies are *tier-based strictly local*: a tier is defined as the projection of a subset of the segments of the input string, and the grammar constraints act only over that subset. For instance, the example below (from AARI, an Omotic language of south Ethiopia) shows how to enforce long-distance sibilant harmony (in anteriority) by projecting a tier T only containing sibilants from the input, and ban contiguous *ʒs and *sʒ on T.

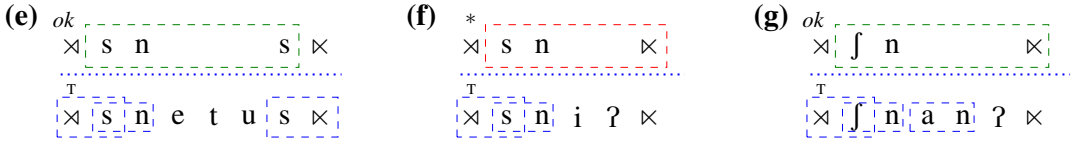
Limitations of TSL TSL grammars project segments on a tier by checking whether they belong to a specific *tier-alphabet*. However, there are processes for which it seems that a TSL grammar needs to access additional structural information that is available in the input string, in order to decide whether to project a segment on a tier. Consider the case of sibilant harmony in the SAMALA language of Southern California, in which a regressive sibilant harmony with unbounded locality ([*s*] and [ʃ] may not co-occur anywhere within the same word, cf. (a)) overrides a restriction against string-adjacent **sn* that results in a pattern of dissimilation (cf. (Applegate 1972)). For example /*sn*/ surfaces as [ʃn] (cf. (b) vs (c)), unless there is an [*s*] following in the string, in which case it surfaces as [*sn*] (cf. (d)).



McMullin (2016) argues extensively that it is not possible to capture this overall pattern with a single TSL grammar. Since [*sn*] is sometimes observed in a string-adjacent context (as in (d)), it must be permitted as a 2-factor on a tier (even though it is only allowed when a segment such as [*s*] follows them later in the string). But then, a TSL grammar would have no means of banning **sn* when there is no subsequent [*s*] in the string (as in (b) vs. (c)). A careful reader might point out that the difference between (c) and (d) can be resolved by extending the tier-grammar to consider 3-factors. However, the grammar needs to ban all instances of **sn* and thus, every occurrence of [*n*] in the input string must projected on the tier. Since the number of n segments between two sibilants is potentially unbounded, no TSL grammar can generally account for this pattern, independently of the dimension of the k-factors on the tier.

Proceedings of the Society for Computation in Linguistics (SCiL) 2018, pages 213-214.
Salt Lake City, Utah, January 4-7, 2018

Structure-Sensitive TSL This kind of expressivity can be accomplished by increasing the locality window of the *tier projection mechanism*. For example, by increasing the locality of the projection to 2, we can allow the grammar to project $[n]$ iff it is immediately preceded by a sibilant in the input string, and then use 3-local tier constraints to ban $\{{}^{*}sn(\neg s), {}^{*}\textstyle\int ns\}$, in addition to the factors needed to enforce the usual sibilant harmony patterns.



This time, the possible unboundedness of n is not a problem, since these segments are now relevant for the projection only when adjacent to a sibilant (cf. (g)). Whereas the TSL projection mechanism is based on each segment's individual (also, 1-local) properties, this new class of grammars — *Structure-Sensitive TSL (SS-TLS)* — relies on structural relationships between segments in the input string (i.e. adjacency to other elements) to choose what to project. Other processes covered by this extension are, for example, the nasal harmony in YAKA — where $[+nasal]$ segments start an harmony domain only if they are not immediately followed by voiced oral stops (Walker 2000) — the culminativity patterns of Heinz (2014), and even the unbounded tone plateauing processes discussed from a computational perspective by Jardine (2016).

Discussion By extending the locality of TSL projections, I show how standard TSL grammars can be viewed as Input Strictly 1-Local mappings — ISL, as in Chandlee (2014)— combined with a Strictly Local (SL) filter applied to their output. Thus, the increase in descriptive power brought by SS-TSL comes with almost no additional formal cost, since k-ISL functions and n-SL grammars are independently used in phonology. Furthermore, SS-TSL languages are clearly *Gold learnable*. It is known that TSL languages are learnable in the limit from positive data since, given a fixed alphabet and a fixed k, the number of possible tiers and permissible tier k-factors is finite. This result extends to SS-TSL languages, which are finite given upper bounds on their fundamental parameters (locality of constraints *and* locality of projection). Besides general learnability, proving that SS-TSL is learnable with a computationally efficient method has important consequences for the cognitive relevance of the TSL neighborhood. Interestingly, the 2-TSL learner proposed by Jardine and Heinz (2016) seems to be easily adaptable to SS-TSL, by inducing a tier of segments based on their k-local properties.

Conclusion A growing body of literature is exploring TSL languages as a good computational hypothesis for the complexity of phonotactic patterns. However, the TSL class comes with particularly tight constraints on the projection function. Here I relax some of these constraints, allowing for a more general definition of *tier-projection*. The resulting new class naturally extends TSL while preserving all its formal properties, and easily captures patterns in which local and non-local dependencies interact, that have been a problem for TSL accounts. This result supports a generalized analysis of phonotactic patterns in terms of subregular complexity. It also shows how understanding the way constraints on the projection mechanism restrain the generative power of TSL could help identify fundamental properties of phonotactic dependencies.

Applegate, R. B. 1972. *Ineseno Chumash grammar*. Ph.D. thesis, U. of California, Berkeley. • **Chandlee J**. 2014. *Strictly Local Phonological Processes*. Ph.D. thesis, U. of Delaware. • **Heinz, J**. 2014. Culminativity times harmony equals unbounded stress. In *Word Stress: Theoretical and Typological Issues*, Chap 8. • **Heinz, J., Rawal C., & Tanner, H. G.** 2011. Tier-based strictly local constraints in phonology. *Proceedings of ACL 49th*, 58–64. • **Jardine, A.** 2016. Computationally, tone is different. *Phonology*, 33(2), 247-283. • **Jardine, A. & Heinz, J**. 2016. Learning tier-based strictly 2-local languages. *Transactions of the ACL*, 4:87–98. • **McMullin, K**. 2016. *Tier-based locality in long-distance phonotactics: learnability and typology*. Ph.D. thesis, U. of British Columbia. • **Walker, R**. 2000. Yaka nasal harmony: Spreading or segmental correspondence? *Annual Meeting of the BLS*, 26(1):321–332.