Tiers and Relativized Locality Across Language Modules

Keywords: computational linguistics, subregular complexity, locality, tiers

Overview Heinz and Idsardi (2013) draw attention to a profound computational difference between syntax and phonology: phonology only requires regular computations over strings (Johnson 1972; Kaplan and Kay 1994), whereas syntax involves non-regular computations over strings (Chomsky 1956; Huybregts 1984; Shieber 1985; Radzinski 1991; Michaelis and Kracht 1997; Kobele 2006). We offer an alternative picture that is synthesized from several recent works in computational linguistics and closes the apparent chasm between phonology and syntax.
The computational complexity of a linguistic domain can be measured along two axes: the nature of the structural representations, and the power of the computations that manipulate these structures. The previously observed complexity differences between syntax and phonology (and morphology) can be recast entirely in terms of the data structures rather than the power of the computations. Phonology operates over string-like structures, whereas syntax uses trees of unbounded size, which grant it increased expressivity. But the dependencies we find in both involve similar computations that fall into the formal class *tier-based strictly local*. Among other things, this entails that the linguistic notion of relativized locality plays a crucial role across language modules.

Measuring Complexity The *subregular hierarchy* (Rogers et al. 2013) provides a very flexible measuring rod for estimating the computational demands of linguistic dependencies. It distinguishes numerous classes, which differ in how sophisticated the grammatical inference rules are, how much memory is needed, and what can be stored in it. This provides a close link between subregular complexity and cognition (Finley 2008; Hwangbo 2015; Avcu 2017). Only two classes are of importance to us: the very weak class *strictly local* (SL), and its linguistically natural extension *tier-based strictly local* (TSL; Heinz et al. 2011).
Intuitively, a linguistic dependency is SL-k iff there is some fixed k such that the dependency applies within a domain of k adjacent symbols. Consider the case of intervocalic voicing. It can be construed as a constraint $^*\mathrm{V}[\mathrm{C},-\mathrm{voiced}]\mathrm{V}$ on surface forms (we put aside mappings from underlying representations to surface forms and focus on directly regulating the shape of surface forms). Intervocalic voicing is SL-3 because it suffices to check for any given string that it does not contain three adjacent symbols s_1, s_2, and s_3 such that s_1 and s_3 are vowels and s_2 is a voiceless consonant.
A dependency is TSL-k iff it can be made SL-k by ignoring certain symbols in the string. Consider unbounded sibilant harmony, which requires all sibilants in a word to agree in anteriority, no matter how far apart they are. This phenomenon is not SL since no matter what k we pick, there will be some string where two sibilants are separated by more than k segments (we make the standard assumption that there is no upper bound on the length of possible words). But unbounded sibilant harmony is TSL-2: we ignore all non-sibilants by constructing a sibilant-tier and forbid any two adjacent symbols on that tier to disagree in anteriority. This is illustrated below with an example from Samala (Applegate 1972):

(1)

s ʃ	ʃ ʃ
*h a s x i n t i l a w a ʃ	h a ʃ x i n t i l a w a ʃ

TSL Phonology and Morphology The dependencies found in phonology are largely TSL (see Heinz 2018, McMullin 2016, and references therein). Some potential counterexamples have been pointed out (Jardine 2016; McMullin and Hansson 2015; McMullin 2016; Baek 2017;

Yang 2018) but have since been shown to fall into TSL if the tier projection function can take a segment's local context into account (De Santo and Graf 2017). Structurally condition tier projection is a linguistically natural addition to TSL, and will be important for the comparison to syntax.
As an example, consider the case of unbounded tone plateauing, where no low tones (L) may occur within an interval spanned by two high tones (H). Hence LHLLL and LLLHL are both well-formed, and so is LHHHHHL, but LHLLLHL is ill-formed. Now suppose that we construct a tier that contains all H irrespective of their structural context, and in addition every L that occurs immediately after an H. Then the SL-3 constraint *HLH will correctly distinguish the well-formed strings from the ill-formed ones. But crucially this presumes that L is not projected in any other configuration, otherwise the two high tones might be arbitrarily far apart.

(2)

H L
L H L L L L L

H L
L L L L L H L

H L H L
*L H L L L H L

Similar studies suggest that morphology is TSL (Aksënova et al. 2016). It seems, then, that phonology and morphology (when restricted to surface forms) can be understood in terms of relativized locality restrictions on strings: tiers are a metaphor for picking out the relevant symbols from the surface string, and only those symbols are subject to local constraints while the rest is ignored.

TSL Syntax TSL can be generalized from strings to trees, and then be applied to formal models of syntax such as Minimalist grammars (MGs; Stabler 1997). In doing so, Graf and Heinz (2015) and Graf (2018b) show that TSL also includes the computational core of syntax — the structure-building operations Merge and Move.
We briefly illustrate this point for Merge in MGs — the argument is very similar for Move. Consider the DP *the girl*, which is built by merging *the* and *girl*. In MGs, each operation must be licensed by features, so *the* must carry a selector feature [Subcat: N] and *girl* must have a matching category feature [Cat: N]. We may view this matching condition as a constraint on syntactic derivations like the ones in (3).

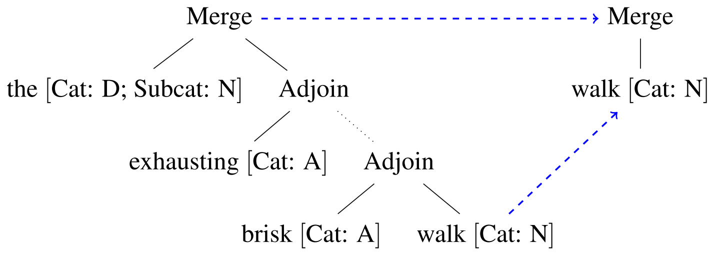
(3) *Well-formed and ill-formed derivation*

Merge
the [Cat: D; Subcat: N] walk [Cat: N]

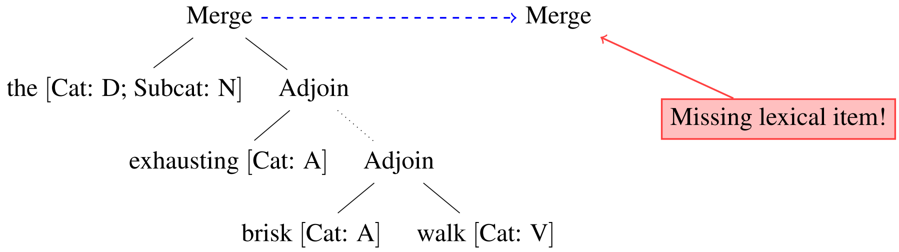
*Merge
the [Cat: D; Subcat: N] walk [Cat: V]

Checking for the presence of matching features is an SL constraint in (3) since the selector and the selectee are in a local configuration. But in fact an unbounded number of NP-adjuncts may occur between the two, turning selection into a long-distance dependency between features. The dependency remains TSL, though. One first constructs an N-tier that contains all lexical items with [Cat: N] and all the Merge nodes that check a [Subcat: N] feature. The Subcat feature checked by a Merge node is always part of its local context, so this kind of tier projection is still TSL. Over such a tier, every Merge node must have exactly one lexical item among its daughter nodes, and every LI must be the daughter of a Merge node.

(4) a. Well-formed derivation with licit N-tier



b. Ill-formed derivation with illicit N-tier



We skip over many technical details here as the important insight is that once again relativized locality is at play: first tier projection identifies the relevant symbols that are subject to a given constraint, and then this constraint applies in a local fashion. As a result, Merge can be applied correctly with TSL-means, similar to numerous processes in phonology.

Cognitive Parallelism and its Implications Previous research has shown that phonology, morphology and syntax look remarkably similar from a subregular perspective. Based on these results we make the following conjecture:

(5) *Strong Cognitive Parallelism Hypothesis*
For every subregular class C, there is a phonological dependency that belongs to C over strings iff there is a syntactic dependency that belongs to C over trees.

In other words, phonology and syntax have the same subregular complexity over their respective data structures. This conjecture allows data from one domain to bear directly on the other. Graf (2018a), for example, uses this to explain the absence of first-last harmony as a specific instance of the principle that local and non-local information in phonology can only interact in a fashion similar to c-command in syntax.

Conclusion Taking as our vantage point the linguistically plausible assumption that phonology operates over string-like structures and syntax over (arbitrarily large) trees, we unearthed a surprising similarity between the two: non-local dependencies between nodes/segments are local within some suitable relativization domain, formalized via tiers. This is not just a simple coding trick such that every non-local dependency is local over a suitable choice of tiers — tier-based formalisms are provably incapable of expressing many (empirically unattested) long-distance dependencies. A linguistically informed choice of data structure thus highlights profound parallels in the computational complexity of phonological and syntactic dependencies, establishing a tight link between these two language modules.

References

Aksënova, Alëna, Thomas Graf, and Sedigheh Moradi. 2016. Morphotactics as tier-based strictly local dependencies. In *Proceedings of the 14th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, 121–130. URL `https://www.aclweb.org/anthology/W/W16/W16-2019.pdf`.

Applegate, Richard B. 1972. *Ineseño Chumash grammar*. Doctoral Dissertation, University of California, Berkeley.

Avcu, Enes. 2017. Experimental investigation of the subregular hierarchy. In *Proceedings of the 35st West Coast Conference on Formal Linguistics at Calgary, Alberta Canada, 2017 April.*. (To appear).

Baek, Hyunah. 2017. Computational representation of unbounded stress: Tiers with structural features. Ms., Stony Brook University; to appear in *Proceedings of CLS 53*.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2:113–124.

De Santo, Aniello, and Thomas Graf. 2017. Structure sensitive tier projection: Applications and formal properties. Ms., Stony Brook University.

Finley, Sara. 2008. *Formal and cognitive restrictions on vowel harmony*. Doctoral Dissertation, John Hopkins University.

Graf, Thomas. 2018a. Locality domains and phonological c-command over strings. To appear in *Proceedings of NELS 2017*.

Graf, Thomas. 2018b. Why movement comes for free once you have adjunction. URL `http://ling.auf.net/lingbuzz/003943`, to appear in *Proceedings of CLS 53*.

Graf, Thomas, and Jeffrey Heinz. 2015. Commonality in disparity: The computational view of syntax and phonology. Slides of a talk given at GLOW 2015, April 18, Paris, France.

Heinz, Jeffrey. 2018. The computational nature of phonological generalizations. In *Phonological typology*, ed. Larry Hyman and Frank Plank, Phonetics and Phonology, chapter 5, 126–195. Mouton De Gruyter.

Heinz, Jeffrey, and William Idsardi. 2013. What complexity differences reveal about domains in language. *Topics in Cognitive Science* 5:111–131.

Heinz, Jeffrey, Chetan Rawal, and Herbert G. Tanner. 2011. Tier-based strictly local constraints in phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, 58–64. URL `http://www.aclweb.org/anthology/P11-2011`.

Huybregts, M. A. C. 1984. The weak adequacy of context-free phrase structure grammar. In *Van periferie naar kern*, ed. Ger J. de Haan, Mieke Trommelen, and Wim Zonneveld, 81–99. Dordrecht: Foris.

Hwangbo, H. J. 2015. Learnability of two vowel harmony patterns with neutral vowels. In *Proceedings of the third Annual Meeting on Phonology (AMP 2015). Oct. 10th, Vancouver, Canada.*.

Jardine, Adam. 2016. Computationally, tone is different. *Phonology* 33:247–283. URL `https://doi.org/10.1017/S0952675716000129`.

Johnson, C. Douglas. 1972. *Formal aspects of phonological description*. The Hague: Mouton.

Kaplan, Ronald M., and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics* 20:331–378. URL `http://www.aclweb.org/anthology/J94-3001.pdf`.

Kobele, Gregory M. 2006. *Generating copies: An investigation into structural identity in language and grammar*. Doctoral Dissertation, UCLA. URL `http://home.uchicago.edu/~gkobele/files/Kobele06GeneratingCopies.pdf`.

McMullin, Kevin. 2016. *Tier-based locality in long-distance phonotactics: Learnability and typology*. Doctoral Dissertation, University of British Columbia.

McMullin, Kevin, and Gunnar Ólafur Hansson. 2015. Long-distance phonotactics as tier-based strictly 2-local languages. In *Proceedings of AMP 2014*.

Michaelis, Jens, and Marcus Kracht. 1997. Semilinearity as a syntactic invariant. In *Logical Aspects of Computational Linguistics*, ed. Christian Retoré, volume 1328 of *Lecture Notes in Artifical Intelligence*, 329–345. Springer. URL `http://dx.doi.org/10.1007/BFb0052165`.

Radzinski, Daniel. 1991. Chinese number names, tree adjoining languages, and mild context sensitivity. *Computational Linguistics* 17:277–300. URL http://ucrel.lancs.ac.uk/acl/J/J91/J91-3002.pdf.

Rogers, James, Jeffrey Heinz, Margaret Fero, Jeremy Hurst, Dakotah Lambert, and Sean Wibel. 2013. Cognitive and sub-regular complexity. In *Formal Grammar*, ed. Glyn Morrill and Mark-Jan Nederhof, volume 8036 of *Lecture Notes in Computer Science*, 90–108. Springer. URL http://dx.doi.org/10.1007/978-3-642-39998-5_6.

Shieber, Stuart M. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8:333–345. URL http://dx.doi.org/10.1007/BF00630917.

Stabler, Edward P. 1997. Derivational Minimalism. In *Logical aspects of computational linguistics*, ed. Christian Retoré, volume 1328 of *Lecture Notes in Computer Science*, 68–95. Berlin: Springer. URL https://doi.org/10.1007/BFb0052152.

Yang, Su Ji. 2018. Subregular complexity in Korean phonotactics. Undergraduate honors thesis, Stony Brook University.