Learning Opaque Mappings

There are several current challenges to a formal theory of how the phonological component of grammar is acquired. This research addresses the particular problem of how opaque mappings can be learned from finitely many examples. The concrete result is that the learning algorithms of Chandlee [3, et seq.] and Jardine et al. [12] are shown to be able to learn (in a well-defined and provably correct sense) the six cases of under- and over-application opacity discussed in Baković's [1] typology of opaque mappings. We also discuss the implications of this result for a theory of phonology, its limitations, and future efforts.

The transformation from input underlying representations to output surface representations [5, 14] can be viewed as a *mapping*. Chandlee [3] establishes that significant classes of phonological processes belong to a small subset of the regular mappings. They are *k-Input Strictly Local* (k-ISL), so called because of their Markovian property that the output at any given point is completely determined by the most recently read k input symbols (cf. [16]). Chandlee [3] and Chandlee et al. [4] show how this property naturally leads to a learning algorithm (ISLFLA) for such processes, a result [12] extends with the algorithm SOSFIA.

The present work extends these results to mappings traditionally described as the *interaction* of different phonological processes. In particular we show the opaque interactions described in Baković (2007)'s typology are also k-ISL and therefore learnable by the ISLFLA and SOSFIA algorithms. Take, for instance, Baković's example from Bedouin Arabic, summarized in (1). The result of the interaction of Rule (1a) and (1b) is not surface true, because non-final [i]s survive in the ouput. However, this interaction is ISL when viewed as a single mapping, as in (2) (other vowels and consonants are represented as Cs and Vs). The opaque examples are that /iCV/ (with /i/ being non-final), gets mapped to [CV], whereas /aCV/ is mapped to [iCV]. Both changes are decided on 3-segment input sequences, /iCV/ and /aCV/, respectively, and so the Bedouin mapping is ISL for $k = 3$.

These results challenge the view of phonology as optimization. It is well-known that classical OT and its variants such as Harmonic Grammar (HG) and Harmonic Serialism (HS) cannot represent all opaque mappings [11, 1, 2]. In addition, OT also overgenerates: even with simple markedness and faithfulness constraints, OT predicts non-regular mappings [2, 6, 15] despite the fact that available evidence suggests all phonological mappings are regular [13, 9, 10]. In contrast, the explicit hypothesis that possible phonological mappings are determined by computational properties like k-ISL provides a better approximation to the attested typology with regard to these issues. However, like OT (and HG, HS), the current proposal does not distinguish individual processes from interacting ones.

There are limitations to the current approach. First, like other algorithms such as Recursive Constraint Demotion, it assumes the underlying forms are given, and not learned [c.f. 17]. Second, the theoretical result cannot be applied directly to natural language data because it is implicational: *if* the learning sample contains the following information *then* the mapping is learned. We expect, however, that properly integrating features and natural classes will help the algorithms find the right kind of information in natural language data [c.f. 7]. Third, long-distance processes are not ISL. We expect this can be addressed along the lines of [8], where long-distance phenomena are treated distinctly from local ones.

Despite these limitations, this research shows that a restrictive, learnable theory of phonology that also can represent, contra OT and its variants, opaque mappings is viable. This is why we believe this research represents an important step forward in understanding the nature of the phonological component of grammar and how humans come to learn it.

(1) Counterfeeding-on-focus in Bedouin Arabic [1]

		/katab/	
a.	$i \rightarrow \emptyset$ / __σ	—	
b.	$a \rightarrow i$ / __σ	kitab	(*ktab)

(2) Examples of the composition of rules (a) and (b) from (1) as a mapping

…iC#	$\mapsto$	…iC#	…iCV…	$\mapsto$	…CV…
…aC#	$\mapsto$	…aC#	…aCV…	$\mapsto$	…iCV…

References

[1] Eric Baković. A revised typology of opaque generalisations. *Phonology*, 24(2):217–259, 2007.

[2] Brian Buccola. On the expressivity of Optimality Theory versus ordered rewrite rules. In *Proceedings of Formal Grammar 2012 and 2013*, volume 8036 of *Lecture Notes in Computer Science*, pages 142–158. Berlin: Springer-Verlag, 2013.

[3] Jane Chandlee. *Strictly Local Phonological Processes*. PhD thesis, U. of Delaware, 2014.

[4] Jane Chandlee, Rémi Eyraud, and Jeffrey Heinz. Learning strictly local subsequential functions. to appear in *Transactions of the Association for Computational Linguistics*.

[5] Noam Chomsky and Morris Halle. *The Sound Pattern of English*. Harper & Row, 1968.

[6] Dale Gerdemann and Mans Hulden. Practical finite state Optimality Theory. In *Proceedings of the 10th International Workshop on FSMNLP*, pages 10–19. ACL, July 2012.

[7] Daniel Gildea and Daniel Jurafsky. Learning bias and phonological-rule induction. *Computational Linguistics*, 24(4), 1996.

[8] Jeffrey Heinz. Learning long-distance phonotactics. *LI*, 41:623–661, 2010.

[9] Jeffrey Heinz and William Idsardi. Sentence and word complexity. *Science*, 333(6040):295–297, July 2011.

[10] Jeffrey Heinz and William Idsardi. What complexity differences reveal about domains in language. *Topics in Cognitive Science*, 5(1):111–131, 2013.

[11] William Idsardi. Clarifying opacity. *The Linguistic Review*, 17:199–218, 2000.

[12] Adam Jardine, Jane Chandlee, Rémi Eyraud, and Jeffrey Heinz. Very efficient learning of structured classes of subsequential functions from positive data. In Alexander Clark, Makoto Kanazawa, and Ryo Yoshinaka, editors, *JMLR: Workshop and Conference Proceedings*, volume 34, pages 94–108, 2014.

[13] Ronald Kaplan and Martin Kay. Regular models of phonological rule systems. *Computational Linguistics*, 20:331–78, 1994.

[14] Alan Prince and Paul Smolensky. *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell Publishing, 2004.

[15] Jason Riggle. *Generation, Recognition, and Learning in Finite State Optimality Theory*. PhD thesis, UCLA, 2004.

[16] James Rogers and Geoffrey Pullum. Aural pattern recognition experiments and the subregular hierarchy. In *Tenth Meeting on Mathematics of Language*. UCLA, 2007.

[17] Bruce Tesar. *Output-Driven Phonology: Theory and Learning*. Cambridge Studies in Linguistics, 2014.