Learning stochastic grammars without stochastic rankings

<u>Introduction</u> Phonological processes that are restricted to certain lexical items typically apply stochastically to novel items. Furthermore, it has been recognized that stochastic grammars reflect statistical generalizations in the lexicon (e.g., Hayes & Londe 2006, Albright & Hayes 2003, among others). Learning a stochastic OT grammar from the lexicon, however, is a problem that has been henceforth unsolved.

I offer a learning algorithm that learns from a list of lexical items, building lexical information into an OT constraint hierarchy. Upon encountering a novel item, the grammar applies stochastically and projects lexical trends onto the novel item. The algorithm is implemented as an open-source Java program. In this work, I demonstrate the workings of the algorithm with a case study of Hebrew plural morphology.

<u>The Hebrew case study</u> Hebrew marks the plural by suffixing *–im* to masculine nouns and *–ot* to feminine nouns. Some masculine nouns exceptionally select the feminine *–ot*, and a majority of those nouns have an [o] in their final syllable. This preference for *–ot* in masculine nouns that end in [o] applies productively to novel nouns, as seen in Berent, Pinker & Shimron (1999).

Simplifying greatly, a morphological constraint, AGREE(gender) requires the masculine *–im* on masculine nouns. The markedness LICENSE(mid) requires that mid vowels be either stressed or auto-segmentally linked to a stressed vowel (see 1 below, and cf. mid vowel licensing in Shona, Beckman 2004). Regular nouns that end in [o] require the ranking AGREE(gender) » LICENSE(mid), as in (2), and exceptional nouns that end in [o] require the opposite ranking (3).

<u>The algorithm</u> The algorithm relies on Biased Constraint Demotion (BCD, Prince & Tesar 1999), augmented by a mechanism of Inconsistency Resolution (Pater to appear). Once the Hebrew learner realizes that conflicting evidence causes BCD to stall (4), the learner creates two lexically-specific clones of LICENSE(mid): one clone specific to a lexical item that requires AGREE(gender) » LICENSE(mid), and another clone specific to a lexical item that requires the opposite ranking (5).

From this point on, when the learner hears a new plural form, they associate the new lexical item with one of the existing clones of LICENSE(mid), simply by trying both possible associations and seeing which one generates the observed form. As more words are learned, the number of nouns that take each suffix is built into the grammar.

Words that don't have [o] in their final syllable will produce the same result with either association (6), and will therefore not be associated with either clone. Thus, only words that end in [o] will be listed by the clones of LICENSE(mid).

The resulting grammar is categorical, since all constraints are ranked just like in classical OT (Prince & Smolensky 1993). When a novel singular is encountered, however, clones compete for influence. The novel form is tried out with both clones of LICENSE(mid), and the likelihood of each clone to be selected is directly proportional to the number of lexical items that are already associated with it, creating a stochastic effect based on the categorical grammar.

<u>Competing accounts</u> Unlike the Gradual Learning Algorithm (GLA, Boersma 1997), my algorithm produces a grammar that is stochastic only relative to novel forms; the GLA does not distinguish listed words from novel words. Moreover, the GLA cannot use a lexicon to learn a stochastic grammar, as pointed out in Hayes & Londe (2006).

The USELISTED constraint of Zuraw (2000) distinguishes existing items from novel ones, but unlike my model, it doesn't derive the patterning of novel items from the trend created by the listed items. Thus, my algorithm and its Java implementation are a step forward in accounting for lexicon-based stochastic phenomena within the OT framework.

(1)		Singular	Plural	
	Regular	*alón*	*alon-ím*	‘oak tree’
			[–hi][+hi]	
	Irregular	*xalón*	*xalon-ót*	‘window’
			[–hi]	

(2)

$/\text{alon}_{\text{MASC}} + \{\text{im}_{\text{MASC}}, \text{ot}_{\text{FEM}}\}/$	AGREE(gender)	LICENSE(mid)
☞ a. alon-ím		*
b. alon-ót	*!	

(3)

$/\text{xalon}_{\text{MASC}} + \{\text{im}_{\text{MASC}}, \text{ot}_{\text{FEM}}\}/$	LICENSE(mid)	AGREE(gender)
a. xalon-ím	*!	
☞ b. xalon-ót		*

(4)

	LICENSE(mid)	AGREE(gender)
xalon-ót ~ xalon-ím	W	L
alon-ím ~ alon-ót	L	W

(5)

	LICENSE(mid)$_{\{xalon\}}$	AGREE(gender)	LICENSE(mid)$_{\{alon\}}$
xalon-ót ~ xalon-ím	W	L	
alon-ím ~ alon-ót		W	L

(6)

$/\text{axbar}_{\text{MASC}} + \{\text{im}_{\text{MASC}}, \text{ot}_{\text{FEM}}\}/$	LICENSE(mid) {*xalon, …, axbar*}	AGREE(gender)	LICENSE(mid) {*alon, …, axbar*}
☞ a. axbar-ím			
b. axbar-ót		*	

Selected References

Albright, Adam and Bruce Hayes (2003) *Learning nonlocal environments*. Talk given at the LSA yearly meeting in Atlanta, Jan 4.

Beckman, Jill (2004) Positional Faithfulness, in *Optimality Theory in Phonology: A Reader*, John J. McCarthy (ed), 310–342.

Berent, Iris, Steven Pinker & Joseph Shimron (1999). Default nominal inflection in Hebrew: Evidence for mental variables. *Cognition* **72**. 1–44

Boersma, Paul (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* **21**. 43–58.

Hayes, Bruce and Zsuzsa Londe (2006) Stochastic phonological knowledge: the case of Hungarian vowel harmony. *Phonology* **23**. 59–104

Pater, Joe (to appear) Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In Steve Parker, (ed.) *Phonological Argumentation: Essays on Evidence and Motivation*. London: Equinox.

Prince, Alan and Bruce Tesar (1999) *Learning Phonotactic Distributions*. ROA-353.

Zuraw, Kie (2000). *Patterned Exceptions in Phonology*. Ph.D. dissertation, UCLA.