Michael Becker, UMass Amherst
michael@linguist.umass.edu

82nd LSA Annual Meeting, Chicago, IL
January 3–6, 2008

The role of markedness constraints in learning lexical trends*

- Speakers replicate probabilistic phonological generalizations about their lexicon, i.e. speakers project a grammar from the words of their language.
- I present the case of plural allomorph selection in Hebrew, and show how markedness constraints shape the kinds of generalizations that speakers can extract from their lexicon.
- I propose a learning algorithm that projects an OT grammar from a lexicon. The resulting grammar is deterministic for known words, but applies probabilistically to novel words.

1 Background

Speakers' treatment of novel items reflects trends in the lexicon (Hayes & Londe 2006, Albright & Hayes 2003, Zuraw 2000, and many others). However, speakers can't learn just any trend in their lexicon – they are limited by UG to notice only phonologically plausible patterns, ignoring what Andrew Nevins calls the "surfeit of the stimulus" (Becker, Ketrez & Nevins 2007).

Previous work on lexical trends in Optimality Theory (Prince & Smolensky 1993/2004) faces one of two problems:

(1) With the GLA (Boersma 1997), existing words cannot be distinguished from novel words, since the grammar is probabilistic with respect to all words.

(2) With the USELISTED approach (Zuraw 2000), existing items are distinguished from novel ones, but the patterning of novel items is not derived from the trend created by the listed items.

*I am grateful to Ram Frost of the Hebrew University for his generous help with running the experiment reported here. Thanks to Adam Albright, Outi Bat-El, Lyn Frazier, John McCarthy, Noam Faust, and Joe Pater for wonderfully helpful comments and suggestions. Any remaining errors couldn't possibly be mine.

2 Case study: Hebrew plurals

Hebrew has two plural markers:

–im on most masculine nouns

–ot on most feminine nouns

Most of the masculine nouns that exceptionally take *–ot* have [o] in their final syllable.

The preference for *–ot* in masculine nouns that end in [o] applies productively to novel nouns, as seen in Berent et al. (1999) and in §3 below.

2.1 The trends in the lexicon

Data from Bolozky & Becker (2006). Native (underlyingly stressless) masculine nouns:

(3)

Final vowel	n	*ot*-takers	
u	1101	6	0.5%
i	464	8	1.7%
a	1349	39	2.9%
e	977	31	3.2%
o	523	146	27.9%

The effect of a root [o] is also felt at a distance,

(4)

Vowel pattern	n	*ot*-takers	
aa	589	12	2.0%
oa	102	12	11.8%
ao	163	34	20.9%

... but only when the intervener is [a].

(5)

Vowel pattern	n	*ot*-takers	
oa	102	12	11.8%
oe	288	0	0%
oi	18	0	0%
ou	1	0	0%

2.2 Analysis

Regular nouns allow [o] freely; irregular nouns want an unstressed [o] to be licensed by a near-by stressed [o].

(6)

	Singular	Plural	
		[–hi] [+hi]	
Regular	alón	a l o n - ím	‘oak tree’
		[–hi]	
Irregular	xalón	x a l o n - ó t	‘window’
		[–hi]	
	olám	o l a m - ó t	‘world’

A morphological constraint, ϕ-MATCH, requires the masculine *–im* on masculine nouns. ϕ-MATCH conflicts with [o]-licensing:

(7)

alon$_{\text{MASC}}$ + {im$_{\text{MASC}}$, ot$_{\text{FEM}}$}	ϕ-MATCH	LOCAL(o)
a. ☞ alon-ím		*
b. alon-ót	*!	

Irregular nouns require a high-ranking LOCAL(o):

(8)

xalon$_{\text{MASC}}$ + {im$_{\text{MASC}}$, ot$_{\text{FEM}}$}	LOCAL(o)	ϕ-MATCH
a. xalon-ím	*!	
b. ☞ xalon-ót		*

When the root [o] is farther away from the stressed syllable, it is weakly licensed:

(9)

olam$_{\text{MASC}}$ + {im$_{\text{MASC}}$, ot$_{\text{FEM}}$}	DISTAL(o)	LOCAL(o)	ϕ-MATCH
a. olam-ím	*!	*	
b. ☞ olam-ót		*	*

(10) LOCAL(o)
An [o] must be licensed by virtue of being stressed, or by virtue of being autosegmentally associated to an adjacent stressed [o]

(11) DISTAL(o)
An [o] must be licensed by virtue of being stressed, or by virtue of being autosegmentally associated to some stressed [o]

Ot-takers that don’t have [o] in them are marked with a lexically-specific ranking of a constraint that doesn’t depend on the root vowel, e.g. *ó/HIGH.

(12)

ʃem$_{\text{MASC}}$ + {im$_{\text{MASC}}$, ot$_{\text{FEM}}$}	*ó/HIGH	ϕ-MATCH
a. ʃem-ím	*!	
b. ☞ ʃem-ót		*

Similar requirements on the licensing of [o] are seen in Shona (Beckman 1997; Hayes & Wilson to appear).

3 Speakers’ knowledge

I tested speakers’ choice of plural suffix with four vowel patterns: [aa], [oa], [ao], and [io].

3.1 Materials and methods

The participants were given novel words presented as masculine nouns, e.g.:

(13) ze *kamoz*, ve-ze od *kamoz*. beyaxad, ele ʃney _____
This$_{\text{MASC}}$ is a *kamoz*, and this$_{\text{MASC}}$ is another *kamoz*. Together, they’re two$_{\text{MASC}}$ _____

Then, the participants heard two possible plurals, e.g. *kmoz-ím* and *kmoz-ót*, and were asked to choose the form that sounded most appropriate. The stimuli are listed in appendix A.

The participants were 53 adult native speakers of Hebrew, students at the Hebrew University in Jerusalem. 10 additional subjects were excluded for making more than 8% mistakes with the actual words tested (see appendix A).

3.2 Results

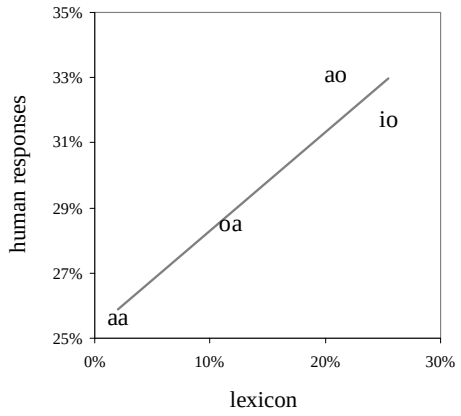
Subjects chose *–ot* least often with [aa], more often with [oa], and most often with [ao], replicating the lexical trend. They did not replicate the lexical difference between [io] and [ao].

(14)

Vowel pattern	Experiment	Lexicon
aa	25.7%	2.0%
oa	28.6%	11.8%
ao	33.2%	20.9%
io	31.8%	25.5%

While the absolute numbers are scaled differently in the lexicon and in the experimental results, they correlate remarkably well:

(15)



The vowel effect came out statistically significant (ANOVA: $F(3,50) = 3.723$, $p = .017$).

This result is in exact accordance with results from Hayes & Londe (2006), who find the same kind of effect in exceptions to vowel harmony in Hungarian.

4 Learning lexical trends

Humans' ability to learn trends in their lexicon is limited by the inventory of possible markedness constraints available. Trends that are beyond the purview of plausible markedness constraints are ignored.

4.1 Ranking conflicts trigger the formation of generalizations

When lexical items demand conflicting rankings, BCD (Prince & Tesar 1999) detects inconsistency and stalls:

(16)

	LOCAL(o)	ϕ-MATCH
a. xalon-ót ≻ *xalon-ím	W	L
b. alon-ím ≻ *alon-ót	L	W

Pater's (2006) solution: Clone a constraint to resolve the inconsistency. My proposal: Make both clones lexically specific.

(17)

	LOCAL(o)$_{xalon}$	ϕ-MATCH	LOCAL(o)$_{alon}$
a. xalon-ót ≻ *xalon-ím	W	L	
b. alon-ím ≻ *alon-ót		W	L

The result: A categorical grammar for listed lexical items.

(18) LOCAL(o)$_{\{xalon,\ makom,\ ...\}}$ ≫ ϕ-MATCH ≫ LOCAL(o)$_{\{alon,\ ʃaon,\ pagoʃ,\ ...\}}$

One clone of LOCAL(o) collects *ot*-takes with a final [o], and the other clone collects *im*-takes with a final [o]. Vowels earlier in the word are ignored, so [io] words and [ao] words get lumped together.

(19) LOCAL(o)$_{146\ items}$ ≫ ϕ-MATCH ≫ LOCAL(o)$_{377\ items}$

The relative number of lexical items on each clone defines a stochastic grammar, which can then apply to novel items.

(20) LOCAL(o)$_{27.9\%}$ ≫ ϕ-MATCH ≫ LOCAL(o)$_{72.1\%}$

4.2 Learning specific patterns

Speakers learn detailed lexical trends by cloning specific constraints first.

(21)

	LOCAL(o)	DISTAL(o)	*ó/HIGH	ϕ-MATCH
a. xalon-ót ≻ *xalon-ím	W	W	W	L
b. olam-ót ≻ *olam-ím		W	W	L
c. ʃem-ót ≻ *ʃem-ím			W	L
d. alon-ím ≻ *alon-ót	L	L	L	W
e. olar-ím ≻ *olar-ót		L	L	W
f. ʃed-ím ≻ *ʃed-ót			L	W

DISTAL(o) prefers *ot*-takers with [o] anywhere in the word.
If a speaker mistakenly clones DISTAL(o) first, one of its clones will gather all the *ot*-takers that have [o] in them, missing the fact that *ot*-takers with a final [o] are more common than *ot*-takers with a non-final [o].

(22) ☹ DISTAL(o)$_{\{xalon,\ olam\}}$ ≫ ϕ-MATCH ≫ DISTAL(o)$_{\{alon,\ olar\}}$

Cloning LOCAL(o) first takes care of final-[o] words, leaving the non-final-[o] words for the care of DISTAL(o):

(23) ☺ LOCAL(o)$_{\{xalon\}}$, DISTAL(o)$_{\{olam\}}$ ≫ ϕ-MATCH ≫ LOCAL(o)$_{\{alon\}}$, DISTAL(o)$_{\{olar\}}$

The most general constraint, *ó/HIGH, is cloned last, and *ot*-takers that don't have [o] in them are grouped together.

(24) LOCAL(o)$_{\{xalon\}}$, DISTAL(o)$_{\{olam\}}$, *ó/HIGH$_{\{ʃem\}}$
≫ ϕ-MATCH ≫
LOCAL(o)$_{\{alon\}}$, DISTAL(o)$_{\{olar\}}$, *ó/HIGH$_{\{ʃed\}}$

5 Conclusions

- The case of the Hebrew plurals joins the growing body of work on the role of phonologically plausible generalizations in learning lexical trends. Hebrew speakers learn the distribution of exceptional *ot*-taking roots in terms of their phonological need for licensing a root [o].
- I proposed a learning algorithm that uses constraint cloning to project an OT grammar that applies deterministically to known items, and projects the behavior of known items probabilistically onto novel items.
- The learning algorithm capitalizes on the inventory of markedness constraints that UG offers to learn phonologically plausible trends and ignore implausible ones.

References

Albright, Adam & Bruce Hayes (2003). Rules vs. Analogy in English past tenses: a computational/experimental study. *Cognition* **90:**. 119–161.

Becker, Michael, Nihan Ketrez & Andrew Nevins (2007). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish devoicing neutralization. Ms. UMass Amherst.

Beckman, Jill (1997). Positional faithfulness, positional neutralisation and shona vowel harmony. *Phonology* **14**. 1–46.

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. In *IFA Proceedings*.

Bolozky, Shmuel & Michael Becker (2006). Living Lexicon of Hebrew Nouns. Ms. UMass Amherst.

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

Hayes, Bruce & Colin Wilson (to appear). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* .

Pater, Joe (2006). The locus of exceptionality: Morpheme-specific phonology as constraint indexation. In Adam Werle Leah Bateman (ed.) *UMOP: Papers in Optimality Theory III*, Amherst, MA: GLSA. 1–36.

Prince, Alan & Paul Smolensky (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Oxford: Blackwell. [ROA-537].

Prince, Alan & Bruce Tesar (1999). Learning phonotactic distributions. Technical Report RuCCS-TR-54, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick [Available as ROA-353].

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

A Experimental materials and results

Nonce words and the percent of *ot*-plurals chosen for them:

(25)

aa		ao		io		oa	
sagaf	9%	zarof	25%	idof	26%	donaf	30%
takav	23%	davov	32%	xizov	25%	ʃolav	25%
kalam	32%	gaʃom	32%	dimom	21%	sotam	38%
garad	38%	kanod	55%	nidod	53%	opad	26%
pasas	34%	baros	23%	migos	25%	xodas	19%
gavaz	9%	kamoz	38%	rizoz	49%	nokaz	21%
banac	21%	pacoc	40%	lixoc	43%	motac	38%
dalaʃ	28%	tanoʃ	32%	biyoʃ	28%	rokaʃ	26%
paʃaʃ	43%	bakoʃ	23%	giroʃ	13%	kovaʃ	13%
zavak	17%	sakok	32%	ʃibok	11%	losak	42%
cagag	38%	barog	30%	ricog	30%	ʃonag	28%
bazax	21%	ʃadoax	47%	lifoax	40%	sovax	21%
ʃanal	28%	calol	25%	zihol	32%	gomal	28%
dagar	19%	galor	32%	cikor	49%	zovar	45%

In the plural forms, the initial [a] was deleted for [aa] and [ao] nouns, e.g. the plurals offered for *sagaf* were *sgaf-im* and *sgaf-ot*.

Real words:

(26)

aa	ao	io	oa
xaʃaʃ (-ot)	makor (mekor-ot)	cinor (-ot)	olam (-ot)
zanav (znav-ot)	xalom (-ot)	nixoax (nixox-ot)	mosad (-ot)
mazal (-ot)	garon (gron-ot)	vilon (-ot)	ocar (-ot)
nahar (nehar-ot)	ason (-ot)	kinor (-ot)	morad (-ot)
davar (dvar-im)	alon (-im)	kidon (-im)	gozal (-im)
bacal (bcal-im)	ʃaon (ʃeon-im)	kiyor (-im)	kolav (-im)

The plural form that was assumed to be correct is in parentheses, with the full form given if it differs from the simple concatenation of the singular and the plural suffix. Subjects who chose the wrong suffix for more than one of these words were excluded.