
What do Hebrew speakers know about irregular patterns? The role of Universal Grammar

- In a study of irregular plurals in Hebrew, I show that lexical exceptions follow phonological patterns. My experimental results show that speakers are aware of the patterns, and extend them productively to novel forms.
- I offer an analysis of Hebrew plurals in terms of universal phonological constraints on plural forms.
- In an artificial language experiment, I show that Hebrew speakers use constraints on surface forms to learn patterns of affix selection (*-im* vs. *-ot*), even in the absence of evidence from real Hebrew.

1 The Hebrew lexicon

1.1 What is irregular morphology?

Regular morphology applies in a predictable way: given any base form, the product is completely known in advance. Examples:

- (1) English progressive: add *-ing*.
“I like to *talk*. In fact, I am *talk-ing* right now.”
- (2) Hebrew definite article: add *a-*
“*ze sefer*. *jel mi a-sefer a-ze?*”

In irregular morphology, the product is not completely predictable given a base form.

- (3) English past tense: add *-ed*, or change the vowel to *u*, or ...
“I try to *sneak* in. Yesterday, I *sneak-ed* / *snuck* in successfully.”
- (4) Arabic plurals: for a noun in *qatl*, choose a plural from *qita:l*, *qutu:l*, ...
singular *baħr*, plural *biħa:r* / *buħu:r* / *ħabħa:r* / *ħabħur*

- (5) Hebrew plural of masculine nouns: add either *-im* or *-ot*.
singular *fofar*, plural *fofar-im* / *fofar-ot*
- (6) Turkish possessive: add a high vowel, and either voice the stem-final stop or not.
bare noun: *efarp*, possessive: *efarb-i* / *efarp-i* (Becker et al. 2007)

1.2 Hebrew plurals

One can discover that *-im* is masculine and *-ot* is feminine by looking at nouns that take different plural suffixes according to natural gender, and then by the completely regular agreement on adjectives and verbs.

- (7) yelad-ím nexmad-ím far-ím
boy-pl nice-pl sing-pl ‘nice boys are singing’

yelad-ót nexmad-ót far-ót
girl-pl nice-pl sing-pl ‘nice girls are singing’

In the native vocabulary, however, masculine nouns can irregularly take *-ot*, and feminine nouns *-im*. The true gender of the noun is revealed by agreement on adjectives and verbs:

- (8) xalon-ót gdol-ím niftax-ím
window-pl big-pl open-pl ‘big windows are opening’
- (9) nemal-ím ktan-ót nixnas-ót
ant-pl small-pl enter-pl ‘small ants are coming in’

Exceptional selection of the plural suffix is only possible when the suffix gets stressed, as in *xalon-ót* and *nemal-ím*. If the stem keeps the stress (usually in loanwords, see Becker 2003), the plural is selected based on the appearance of the singular, without exception – this is categorical irregularity.

- (10) fukátġ-a * fukátġ-im fukátġ-ot ‘focaccia’
blóg blóg-im * blóg-ot ‘blog’
kolég-a * kolég-im kolég-ot ‘(male) colleague’
madám ? madám-im * madám-ot ‘madam (in a brothel)’

1.3 Patterns of irregularity

For any given noun of Hebrew, the selection of *-im* or *-ot* is unpredictable, and often categorical. There are even some minimal pairs:

- (11) a. tor-ím / tor-ót 'line, queue', 'appointment' / 'turn'
 b. himnon-ím / himnon-ót 'national anthem' / 'religious hymn'
 c. maamad-ím / maamad-ót 'stand' / 'status'

The lexicon as whole, however, reveals patterns of irregularity. Looking at the native masculine nouns in Bolozky & Becker (2006), it is clear that an [o] in the stem correlates with taking *-ot* (see also Glinert 1989; p. 454, Aronoff 1994; p. 76).

(12)

Final vowel	<i>n</i>	<i>ot</i> -takers		
V	3891	84	2.2%	gvúl, safék, yáin, ocár
o	523	146	27.9%	sód, xalóm, vilón
Total	4414	230	5.2%	

The effect of [o] can be further refined: It is also felt at a distance, as seen in the disyllabic nouns in (13).

(13)

Vowel pattern	<i>n</i>	<i>ot</i> -takers		
aa	589	12	2.0%	zanáv, xafád
oa	102	12	11.8%	mosád, gorál
ao	163	34	20.9%	garón, maxóz

Is this correlation between a stem [o] and *-ot* a mere historical accident, or is it useful information for speakers? Do speakers incorporate this correlation into their active, synchronic grammar?

If speakers have this knowledge, how have they learned it?

2 Speakers' knowledge, part I

If speakers learn a pattern that goes beyond memorizing the real words of Hebrew, the pattern can be tested with possible, but unattested words.

I used a novel word task to see what kind of knowledge speakers project from the existing words of their language, or how they generalize/analogize from their lexicon.

2.1 The task

Hebrew speakers were asked to choose a plural suffix for nouns of four vowel patterns:

(14)

[a a]	[o a]	[a o]	[i o]
sagáf	donáf	zaróf	ʔidóf
takáv	ʃoláv	davóv	xizóv
kalám	sotám	gaʃóm	dimóm
garád	ʔopád	kanód	nidód
pasás	xodás	barós	migós
gaváz	nokáz	kamóz	rizóz
banác	motác	pacóc	lixóc
daláf	rokáf	tanóf	biyóf
paʃáf	kováf	bakóf	giróf
zavák	losák	sakók	ʃibók
cagág	ʃonág	baróg	ricóg
bazáx	sováx	ʃadóax	lifóax
ʃanáʃ	gomál	calól	zihól
dagár	zovár	galór	cikór

The participants were given the novel words in random order, presented as masculine nouns, e.g.:

- (15) ze *kamoz*, ve-ze od *kamoz*. beyaxad, ele ʃney _____
 This_{MASC} is a *kamoz*, and this_{MASC} is another *kamoz*. Together, they're two_{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 participants were 53 adult native speakers of Hebrew, students at the Hebrew University in Jerusalem. I am grateful to Ram Frost, of the Hebrew University psychology department, for his generous help with running the experiment.

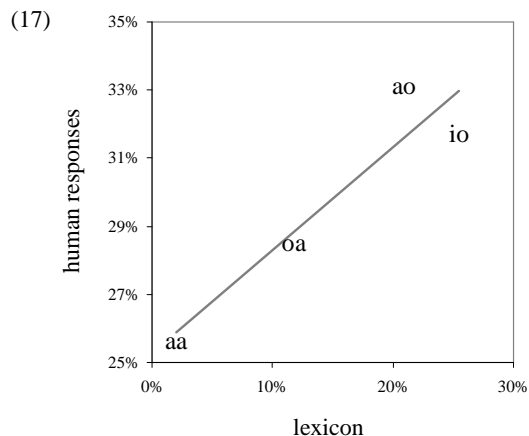
2.2 Results

Speakers chose *-ot* least often with [aa], more often with [oa], and most often with [ao] and [io], replicating the lexical trend.

Since all vowel patterns were equally represented in the experiment, the differences between them are not coming from the experiment, but rather from speakers' brains.

(16)	Lexicon	Experiment
least likely to take <i>-ot</i>	VV	VV
	oV	oV
most likely to take <i>-ot</i>	Vo	Vo

The percentage of *ot*-taking in the lexicon and in the experiment correlate remarkably well:



The only mismatch between the lexicon and the experiment ([ao] vs. [io]) is exactly where we expect speakers not to care – they care about the presence of [o], other vowels don't matter.

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

2.3 Conclusion

Speakers know that having [o] in the stem is conducive to choosing *-ot*, and they know that this effect is stronger when the [o] is in the final syllable. Vowels that precede the stem [o] don't matter.

3 Analysis

3.1 Licensing [o] in world languages

If a sound X is only allowed in some position, the position *licenses* the sound.

Many languages require [o] to be licensed by the stressed syllable:

(18) Russian allows [o] only in the stressed syllable: *dóm-a* 'at home', *dám-áx* 'at homes'.

(19) In most dialect of English, [o] can be unstressed ('piano', 'fellow'), but in some dialects, unstressed [o] is not allowed ('piana', 'fella').

Other languages require [o] to be licensed by the word-initial syllable:

(20) Turkish native nouns allow [o] only in the first syllable of the word.

(21) Shona allows [o] in the word-initial syllable, and an initial [o] can license an [o] later in the word (Beckman 1997; Hayes & Wilson to appear)

Hebrew will turn out to be like Shona, but with stress: In Hebrew, [o] must be stressed, but a stressed [o] allows [o] to appear elsewhere in the word.

3.2 [o]-licensing in Hebrew

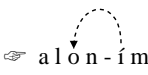
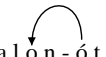
Hebrew plurals are formed in response to conflicting constraints (= requirements, conditions) on surface forms (Optimality Theory, Prince & Smolensky 1993/2004): On the one hand, the demand for matching masculine stems with the masculine *-im*, and on the other hand, the demand for the vowel [o] to be licensed by a stressed [o].

(22) Regular a l ó n a l ó n - í m 'oak tree'

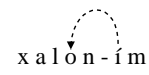

(23) Irregular x a l ó n x a l ó n - ó t 'window'

The need for gender-matching and the need for licensing [o] are in conflict; only one constraint can be satisfied at the expense of the other, if there is an unstressed [o] in the stem.

(24) Taking *-im* to satisfy MATCH GENDER

	MATCH GENDER	LICENSE [o]
a.  a l o n - í m	✓	✗
b.  a l o n - ó t	✗	✓

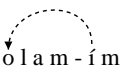
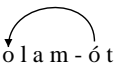
(25) Taking *-ot* to satisfy LICENSE [o]

	LICENSE [o]	MATCH GENDER
a.  x a l o n - í m	✗	✓
b.  x a l o n - ó t	✓	✗

In *olam*, the [o] is unlicensed in the singular, but it gets licensed in the plural:

(26) Irregular  o l á m  o l a m - ó t 'world'

(27) Taking *-ot* to satisfy LICENSE [o]

	LICENSE [o]	MATCH GENDER
a.  o l a m - í m	✗	✓
b.  o l a m - ó t	✓	✗

3.3 Using conflicts to learn a grammar

Hebrew supplies conflicting evidence about the licensing of [o]: in *alon-im*, gender matching is more important than [o]-licensing, but in *xalon-ot*, [o]-licensing is more important than gender matching.

Speakers can detect this conflict (Prince & Tesar 1999), and use it to start keeping track of individual words (Pater 2006; Becker 2006). I implemented this procedure computationally; see Becker (2008) for details.

My approach relies on the idea that exceptions and regular words are responding to different grammatical principles.

4 Using constraints to learn morphology

How do speakers learn that [o] in the stem correlates with *-ot*?

(28) Speakers learn a rule:

“Take a singular with [o], add *-ot* to it to make a plural”

(29) Speakers prioritize conflicting constraints:

“In the plural, licensing an unstressed [o] is better than matching gender”

The rule refers to the singular and to the plural; the constraints refer only to the plural.

In the actual words of Hebrew, every plural stem with [o] always has a singular stem with [o], so the two approaches cannot be distinguished. Theoretically, a Hebrew learner could choose either approach and still get the right result: stem [o] correlating with *-ot*.

If for Hebrew both approaches work, what do speakers do? In the lack of evidence, which approach do humans prefer, i.e. what does Universal Grammar say?

4.1 Speakers' knowledge, part II

I am able to argue in favor of constraints with data from an artificial language experiment. This work was done in collaboration with Lena Fainleib (Tel Aviv University).

4.2 The task

In the artificial languages, singulars were plausible native Hebrew nouns with an [o] or an [i] in their final syllable, and in the corresponding plural forms, the vowels were switched. The choice of the plural suffix agreed with the plural form in the “constraint” language and with the singular form in the “rule” language (30).

(30)	“constraint” language		“rule” language	
	amíg	amog-ót	amíg	amog-ím
	axís	axos-ót	axís	axos-ím
	azíx	azox-ót	azíx	azox-ím
	afív	afov-ót	afív	afov-ím
	adíc	adoc-ót	adíc	adoc-ím
	agóf	agif-ím	agóf	agif-ót
	apóz	apiz-ím	apóz	apiz-ót
	acók	acik-ím	acók	acik-ót
	abóf	abif-ím	abóf	abif-ót
	alód	alid-ím	alód	alid-ót

After speakers learned one of the two languages (the “memorization” stage), they were given nouns in the singular, and were asked to generate the plural (the “generalization” stage).

(31)
agív, apís, axíg, amíx, alíc, axíf, aníz, afík, afíf, azíd
amón, adós, afóg, atóx, afóc, aróf, ahóz, abók, agóf, apód

4.3 The predictions

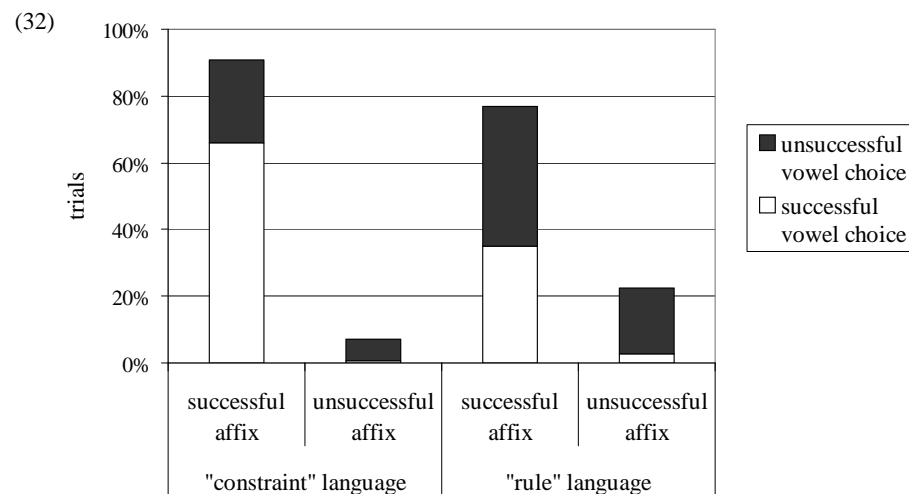
Prediction of my constraint-based approach: Since Hebrew speakers learn that *-ot* licenses an unstressed [o] in the plural stem, it doesn’t matter what the stem vowel is in the singular. Speakers will be more successful in learning the artificial language that pairs *-ot* with stems that have [o] in the plural, i.e. the “constraint” language.

Prediction of the rule-based approach: Speakers learn that singulars with [o] get *-ot* added to them, so they will be more successful in learning the artificial language that pairs *-ot* with stems that have [o] in the singular, i.e. the “rule” language.

4.4 Results

The participants were 26 native speakers of Hebrew – 21 students at the Hebrew University and 5 at the Tel Aviv University. I am grateful to Ram Frost, of the Hebrew University psychology department, for his generous help with running the experiment.

As I predicted, speakers were significantly more successful at generalizing the “constraint” language than the “rule” language ($t(22.74) = 2.23, p = .036$).



The speakers who learned the “constraint” language chose the expected plural affix more often than the speakers who learned the “rule” language, and were more likely to get the stem vowel right.

This result makes sense in light of my theory, in which stem vowels are licensed by the vowel of the plural suffix. When the plural suffix is *-ot*, it’s easy to change a stem vowel to [o], because the [o] would be licensed (in the “constraint” language), and it’s tempting to keep a singular [o] and not change it to [i] (in the “rule” language).

4.5 Summary

In artificial language experiments, it is expected that the more natural language will be easier to learn, everything else being equal.

What do the artificial languages have in common?

- (33)
- The singulars are identical
 - The plural stems are identical
 - The stem-vowel changes are identical ([i] to [o] and [o] to [i])

How do the artificial languages differ? Only in their choice of plural affixes.

How are these languages like Hebrew? Can we be sure that speakers were using their Hebrew grammar when they learned these languages?

- (34)
- Contain nouns that pair sound and meaning
 - Nouns appear in grammatical Hebrew sentences that give them masculine gender, and either singular or plural number
 - Use the vowels and consonants of native Hebrew
 - Use the vocalic patterns and stress patterns (*mishkalim*) of native Hebrew
 - Use *-im* and *-ot* to mark the plural

How are these languages *not* like Hebrew?

- (35)
- Singular [i] changes to plural [o] and vice versa
 - The choice of plural affix is completely predictable from the stem vowel

5 Conclusions

- Hebrew allows masculine nouns to take *-ot* exceptionally. The exceptions follow a pattern: a stem with a final [o] is most likely to take *-ot*, a stem with a non-final [o] is somewhat likely to take *-ot*, and a stem without [o] is least likely to take *-ot*.
- I have shown that speakers identify this pattern in their lexical exceptions, generalize it, and apply it to novel words they encounter. I proposed an analysis of the pattern in terms of constraints on plural forms.

- Results from an artificial language experiment show that Hebrew speakers learn exceptional plurals in terms of constraints on plural forms, even in the absence of evidence for using constraints from real Hebrew.
- I proposed that humans are pre-disposed to learn patterns of irregular morphology in terms of conflicting constraints on output forms, and attributed this pre-disposition to Universal Grammar.

References

- Aronoff, Mark (1994). *Morphology by itself*. MIT Press.
- Becker, Michael (2003). Lexical statification of Hebrew: The disyllabic maximum. In Yehuda N. Falk (ed.) *Proceedings of the Israel Association for Theoretical Linguistics 19*.
- Becker, Michael (2006). Learning the lexicon through the grammar – Evidence from Turkish. Ms. UMass Amherst.
- Becker, Michael (2008). *Phonological Trends in the Lexicon: The Role of Constraints*. Ph.D. dissertation, UMass Amherst.
- Becker, Michael, Nihan Ketzrez & 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.
- Bolozky, Shmuel & Michael Becker (2006). Living Lexicon of Hebrew Nouns. Ms. UMass Amherst.
- Glinert, Lewis (1989). *The grammar of Modern Hebrew*. Cambridge University Press.
- 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 Leah Bateman & Adam Werle (eds.) *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].