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

Why and how to use Universal Grammar in learning irregular morphology[*]

- How do speakers figure out morphology? How do they learn what affixes sound like, and what happens when affixes attach to roots of various shapes? What kind of machinery is involved?

- I will argue that humans use Universal Grammar (UG) to learn their morphology. I will show that Hebrew speakers have a learning bias that isn't coming from Hebrew, and I will attribute it to the structure of human language.

- I will contrast my analysis with a UG-less analysis, and show that it doesn't capture the full range of human behavior.

1 Learning the English past tense

Three ways of pronouncing the past tense suffix:

(1)

[-d]	[-t]	[-ɪd]
[ɹɑb-d]	[stip-t]	[sɑlt-ɪd]
[bʌz-d]	[fʌs-t]	[fold-ɪd]
[pil-d]	[pɪk-t]	[vɛt-ɪd]
[pɔ-d]	[bɹʌʃ-t]	[nid-ɪd]

And speakers know the distribution, of course. Example:

(2)   I love [stɪb]-ing. I [stɪb] every day. Yesterday, I [stɪbd] / [stɪbt] / [stɪbɪd].

But some verbs don't do any of the above. Some verbs stay completely unchanged in the past, like [kʌt] and [ɹɪd]. Do speakers know anything about those?

(3)   I love [stɪb]-ing. I [stɪb] every day. Yesterday, I [stɪbd] / [stɪb].

(4)   I love [snɛd]-ing. I [snɛd] every day. Yesterday, I [snɛdɪd] / [snɛd].

1.1 Learning morphology using UG

Optimality Theory (Prince & Smolensky 1993/2004) is the current mainstream theory in phonology. It works by using one Universal set of *constraints*, which are *ranked* differently in different languages or situations.

Two principles we'll use to analyze English:

(5)   In many languages, words cannot end in two alveolar stops: *dd], *tt], *dt], *td]. There is a universal constraint, *DD], that is always respected in English.

(6)   Adding or removing phonological material is costly, because it masks the underlying representation. There are constraints that penalize change between the UR and the surface.

The English learner hears four different past tense suffixes on the surface: [-d], [-t], [-ɪd] and no change (Ø). Can they find a single underlying representation for the past tense that will work?

(7)   Choosing [-t] as the UR of the past tense:
Works for /pɪk + t/ → [pɪkt]
Works for /bʌz + t/ → [bʌzd] if you allow voicing assimilation
Does it work for /sɑlt + t/ → [sɑltɪd]? Almost certainly not: there is no plausible source for the final [d] in [sɑltɪd].

(8)   Choosing [-ɪd] as the UR of the past tense:
Works for /sɑlt + ɪd/ → [sɑltɪd]
Does it work for /pɪk + ɪd/ → [pɪkt]? Almost certainly not: why would you lose the vowel in [pɪkɪd]?

(9) Choosing [-d] as the UR of the past tense:
Works for /bʌz + d/ → [bʌzd]
Works for /pɪk + d/ → [pɪkt] with voicing assimilation
Works for /sɑlt + d/ → [sɑltɪd]: *DD] forces the addition of a vowel.
Works for /rɪd + d/ → [rɪd]: *DD] forces the deletion of a consonant

Why can't verbs like [stɪb] stay unchanged in the past? /stɪb + d/ → [stɪb] makes no sense: The constraint *DD] doesn't penalize [stɪbd], so nothing forces the deletion of the final [d].

1.2 So what happened to the regular/exceptional dichotomy?

It is largely gone.

(10) The past tense of [stɪb] can only be [stɪbd].
The past tense is predictable, or regular.

(11) The past tense of [snɛd] can be either [snɛdɪd] or [snɛd].
The past tense is not entirely predictable; it allows some exceptionality.

So which verbs are regular? The ones that satisfy *DD] by adding a vowel (salt-ed, fold-ed), or the ones that satisfy *DD] by deleting a consonant (cut, rid)? Both patterns are productive.

What is the past tense of "shed"?

1.3 Learning morphology without UG

The Minimal Generalization Learner (Albright & Hayes 2002, 2003, 2006) learns morphology by creating rules of increasing generality:

(12)

	change	environment
[sɑlt] ∼ [sɑltɪd]	Ø → [ɪd]	/ # s ɑ l t __ #
[bolt] ∼ [boltɪd]	Ø → [ɪd]	/ # b o l t __ #
generalization:	Ø → [ɪd]	/ [+back, −high] l t __ #

(13)

	change	environment
generalization:	Ø → [ɪd]	/ [+back, −high] l t __ #
[niːd] ∼ [niːdɪd]	Ø → [ɪd]	/ # n iː d __ #
generalization:	Ø → [ɪd]	/ [+cor, −ant, −cont] __ #

The resulting generalization misses sometimes: The last rule, which adds [ɪd] before an anterior coronal stop, can apply to [sprɛd] and give *[sprɛdɪd]. That's why each rule has a success rate, defined as its hits divided by its (hits + misses).

(14)

	change	environment
[rɪd] ∼ [rɪd]	no change	/ # r ɪ d #
[kʌt] ∼ [kʌt]	no change	/ # k ʌ t #
generalization:	no change	/ [+cor, −ant, −cont] #

The result: A learner who knows that only {t,d}-final verbs can either take [ɪd] or stay unchanged. Impressive!

This learner could equally well learn a language, English′, in which {t,d}-final verbs take [ɪd] and {p,b}-final verbs stay unchanged. There is no connection between the [d] of the past tense suffix and the reluctance of {t,d}-final verbs to take it.

In fact, the learner works without "knowing" anything about what affixes are or where to look for them. There are no expectations about relationships between the shapes of affixes and the shapes of roots. There are no expectations about what a possible language might be, i.e. no Universal Grammar.

Time to prove that speakers use Universal Grammar to learn morphology!

2 Hebrew plurals

Hebrew has two allomorphs of the plural suffix, [-im] and [-ot].

The learner 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.

(15)   yelad-ím   ktan-ím   ʃar-ím
        boy-pl   little-pl   sing-pl   'little boys are singing'

        yelad-ót   ktan-ót   ʃar-ót
        girl-pl   little-pl   sing-pl   'little girls are singing'

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

(16)   xalon-ót   ktan-ím   niftax-ím
        window-pl   small-pl   open-pl   'little windows are opening'

(17)   cipor-ím   ktan-ót   ʃar-ót
        bird-pl   small-pl   sing-pl   'little birds are singing'

The masculine nouns that take *–ot* are not randomly distributed:

Bolozky & Becker (2006) list 230 masculine nouns that take [-ot], and 146 of them have [o] in their final syllable. Having [o] in the root makes taking [-ot] more likely.

Unsurprisingly, when given a masculine noun they haven't heard before, Hebrew speakers like it better with [-ot] if it has an [o] in it (Berent et al. 1999; Becker 2008).

2.1 Learning Hebrew with Universal Grammar

The learner will identify [-im] as the masculine plural and [-ot] as the feminine plural, but will accept that the two affixes can compete for the same noun, even if its gender is known.

(18)   A universal constraint, MATCH GENDER, wants the masculine [-im] on masculine nouns.

(19)   A universal constraint, LICENSE[o], wants [o] to be licensed by being stressed or by being next to a stressed [o].
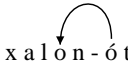
If a sound X is only allowed in some position, the position *licenses* the sound. Many languages require [o] to be licensed by the stressed syllable:

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

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

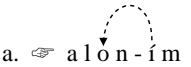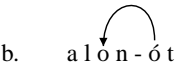Other languages require [o] to be licensed by the word-initial syllable:

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

(23)   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.
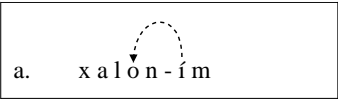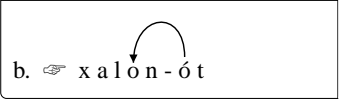
(24)   Regular   a l ó n   a l o n - í m   'oak tree'

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

The constraints MATCH GENDER and LICENSE[o] are in conflict; only one constraint can be satisfied at the expense of the other if there is an unstressed [o] in the root.

(26)   Taking *–im* to satisfy MATCH GENDER

alon$_{\text{MASC}}$ + {im$_{\text{MASC}}$, ot$_{\text{FEM}}$}	MATCH GENDER	LICENSE [o]
a. ☞ a l o n - í m	✔	✗
b. a l o n - ó t	✗	✔

(27) Taking *–ot* to satisfy License [o]

xalon$_{\text{MASC}}$ + {im$_{\text{MASC}}$, ot$_{\text{FEM}}$}	License [o]	Match Gender
a. x a l ó n - í m	✗	✔
b. ☞ x a l o n - ó t	✔	✗

2.2   Learning Hebrew without Universal Grammar

The learner identifies two changes: Ø → [im] and Ø → [ot].
The environments for the two changes are different:

(28)   Ø → [im] has a high success rate with [a], [e], [i], [u], and a somewhat
lower success rate with [o]. But [a, e, i, u] don't make a natural class that
excludes [o], so the general rule is: Ø → [-im] / __ #.

(29)   Ø → [ot] has very low success rate with [a], [e], [i], [u], and a much higher
success rate with [o]. So we get two very general rules:
(a)   Ø → [ot] / __ # (low-ish success rate)
(b)   Ø → [ot] / o C __ # (high success rate)

The learner discovers that having [o] in the root makes choosing [-ot] more likely!

2.3   With or without Universal Grammar?

The UG-less learner takes a singular noun, and decides which affix to attach to it.
If the singular has [o] in it, it is more likely to take [-ot].

The UG-ful learner creates two plural nouns, and chooses which one to keep. If a
plural has an unlicensed [o] in it, it is likely to be rejected.

The UG-less learner makes decisions based on the singular, and the UG-ful learner
makes decisions based on the plural. In real Hebrew, every noun that has [o] in its
plural stem also has [o] in the singular, and almost every noun that has [o] in its
singular stem keeps that [o] in the plural.

Real Hebrew is described equally well by both learners, because in real Hebrew,
the [o] is there both in the singular and in the plural. But what if we made up a
language, Hebrew′, that changes the vowels between the singular and the plural?

3   Artificial Hebrew

This work was done in collaboration with Lena Fainleib (Tel Aviv University), and
benefitted greatly from the generous help of Ram Frost (Hebrew University).

3.1   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
"surface" language and with the singular form in the "deep" language.

(30)

"surface" language		"deep" language	
amíg	amog-ót	amíg	amog-ím
axís	axos-ót	axís	axos-ím
azíx	azox-ót	azíx	azox-ím
aʃív	aʃov-ót	aʃív	aʃov-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
abóʃ	abiʃ-ím	abóʃ	abiʃ-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 "gen-
eralization" stage).

(31)   agív, apís, axíg, amíx, alíc, axíf, aníz, aʃík, afíʃ, azíd
amóv, adós , aʃóg , atóx , aʃóc, aróf , ahóz , abók , agóʃ, apód

7

8

3.2 The predictions

Prediction of my UG-ful approach: When a speaker creates a plural form, the Universal constraint LICENSE[o] wants [o] to be licensed. It doesn't care what the vowel might have been in the singular.

(32) In the "surface" language, LICENSE[o] is always satisfied, so it helps make the right choice of plural suffix

(33) In the "deep" language, LICENSE[o] is violated half of the time, so it dissuades the speaker from making the right choice.

The "surface" language is predicted to be easier to generalize than the "deep" language.

Prediction of the UG-less approach: In real Hebrew, the two available changes are Ø → [im] and Ø → [ot]. In the artificial languages, there are four changes:

(34)

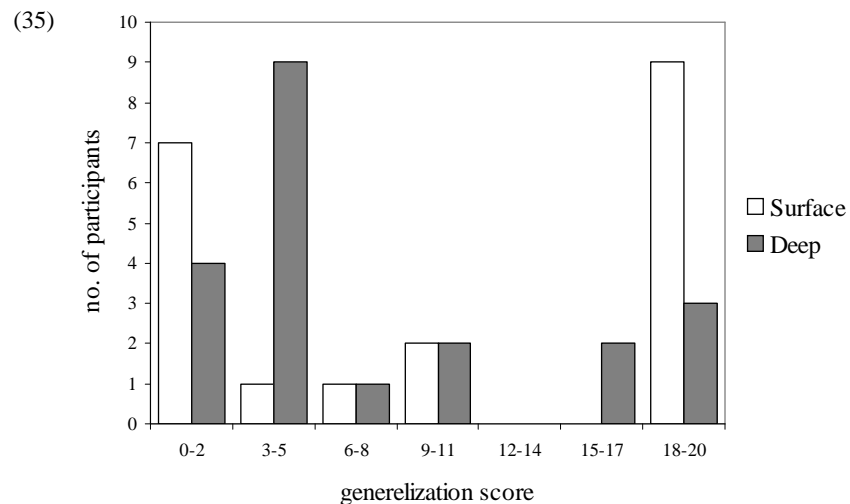| "surface" language	"deep" language
o C → [i C im]	o C → [i C ot]
i C → [o C ot]	i C → [o C im]

The changes of the artificial languages are not found in real Hebrew, and vice versa. Both the "surface" and the "deep" languages are equally distant from Real Hebrew, so they should be equally easy/hard to generalize.

3.3 Results

The participants were 41 native speakers of Hebrew, who were students at the Hebrew University or the Tel Aviv University.

As I predicted, speakers were significantly more successful at generalizing the "surface" language than the "deep" language ($t(39) = 2.16$, $p = .015$).

This difference was not due to the random smartness for the participants who learned the "surface" language, because both groups performed equally well in the memorization stage ($t(39) = .48$, $p > .1$, see appendix A).

(35)

3.4 Summary

How are these languages like Hebrew?

(36) a. Contain nouns that pair sound and meaning

b. Nouns appear in grammatical Hebrew sentences that give them masculine gender, and either singular or plural number

c. Use the vowels and consonants of native Hebrew

d. Use the vocalic patterns and stress patterns of native Hebrew

e. Use [-im] and [-ot] to mark the plural

How are these languages *not* like Hebrew?

(37) a. Singular [i] changes to plural [o] and vice versa

b. The choice of plural affix is completely predictable from the stem vowel

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

(38) Speakers chose [-im] 54% of the time, which is significantly more often than the expected 50% (Wilcoxon test with $\mu=10$, $V = 262$, $p = .027$). This shows that speakers treated the new words as masculine nouns, and used their knowledge about real Hebrew masculine nouns.

4 Conclusions

- I contrasted two approaches to learning morphology: One approach that uses Universal Grammar, and thus has specific expectations about what a human language can be like, and one approach that has no such expectations.

- Both approaches are equally good at learning the plurals of real Hebrew words, stating the connection between having [o] in the root and taking the plural [-ot]. My UG-ful approach states the generalization over plural forms, while the UG-less approach states the generalization over singular forms.

- I asked Hebrew speakers to learn an artificial language that switched [o] and [i] between the singular and plural. Those who learned the artificial language that had [o] going with [-ot] in the plural performed well. Those who learned the artificial language that had [o] going with [-ot] in the singular performed significantly worse.

- Since the preference for stating the generalization over the plurals isn't coming from anything about the real words of Hebrew, I conclude that it is Universal Grammar that's responsible for it.

References

Albright, Adam & Bruce Hayes (2002). Modeling English past tense intuitions with minimal generalization. In Michael Maxwell (ed.) *Proceedings of the sixth meeting of the ACL special interest group in computational phonology*. Philadelphia: ACL.

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

Albright, Adam & Bruce Hayes (2006). Modeling productivity with the gradual learning algorithm: The problem of accidentally exceptionless generalizations. In Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds.) *Gradience in Grammar*, Oxford University Press.

Aronoff, Mark (1994). *Morphology by itself*. MIT Press.

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.

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.

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].

A Participants' memorization scores

(39)