Exceptions

Class 6: Informative URs

Adam Albright

CreteLing 2024

creteling2024.phonology.party

Finding (un)predictability

- A recurring theme of the Kenstowicz and Kisseberth definitions: unpredictable information
  - Contrast = unpredictability
  - Neutralization (absolute, contextual, allophonic) = predictability
- Learning URs = finding unpredictable properties
  - Not including property P $\Rightarrow$ unsuccessful derivation
  - Including property P $\Rightarrow$ successful derivation
  - Successful = accurate
- Limits imposed by set of “formal devices”

Condition B

The UR of a morpheme contains those variant (alternating) and invariant phonetic properties that are idiosyncratic (unpredictable). But it may contain only those variant properties that occur in the PR that appears in isolation (or as close to isolation as the grammar of the language permits).

- Predictability limited by absolute condition
- No learning

Condition B$''$

> The UR of a morpheme may include both variant and invariant phonetic properties. All of the variant properties selected to appear in the UR must occur in a single surface alternant of that morpheme, the basic alternant. The choice of the basic alternant is constrained by a principle of parallelism according to which the basic alternant for all morphemes of a given morphological class (noun, verb, particle, etc.) must occur in the same morphological context.

- Hypothesis space: sets of URs as seen in each paradigm cell
- Sensible objective: choose the set that yields the most accurate grammar (fewest exceptions)

A simple procedure

- Learners are given inflected forms, labeled for stem lexeme morphological features

  gluk ‘car.SG’

  drubi ‘dog.PL’

  zupa ‘cat.SG’

  drup ‘dog.SG’

  gluki ‘car.PL’ (etc.)

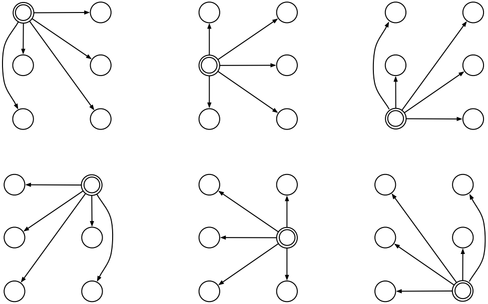
- Morphologically related pairs are assembled
  - *gluk* ∼ *gluki* ‘car-SG/PL’

Schematic input

sap	‘finger.SG’	∼	sapi	‘finger-PL’
drup	‘dog.SG’	∼	drubi	‘dog-PL’
kop	‘friend.SG’	∼	kobi	‘friend-PL’
rat	‘grape.SG’	∼	radi	‘grape-PL’
fet	‘snail.SG.’	∼	feti	‘snail-PL’
tik	‘foot.SG’	∼	tigi	‘foot-PL’
gluk	‘car.SG’	∼	gluki	‘car-PL’
zupa	‘cat.SG’	∼	zupi	‘cat-PL’
kroma	‘chair.SG’	∼	kromi	‘chair-PL’

Searching the space

- Learner tries each inflectional category as input to grammar that derives related forms
  - Grammars are learned using Minimal Generalization Learner, as described in class 1
  - SG $\longrightarrow$ PL, PL $\longrightarrow$ SG (etc.)



Learning: SG $\longrightarrow$ PL direction

$\varnothing \to i$ rules

$\varnothing \to \text{i} \,/\, \text{sap} \,\_\_\,]_{\text{wd}}$

$\varnothing \to \text{i} \,/\, \text{fet} \,\_\_\,]_{\text{wd}}$

$\varnothing \to \text{i} \,/\, \begin{bmatrix} -\text{son} \\ -\text{contin} \\ -\text{dors} \end{bmatrix} \,\_\_\,]_{\text{wd}}$

(2/5: *sap, fet, *drup, *kop, *rat*)

$\varnothing \to \text{i} \,/\, \text{gluk} \,\_\_\,]_{\text{wd}}$

$\varnothing \to \text{i} \,/\, \begin{bmatrix} -\text{son} \\ -\text{contin} \end{bmatrix} \,\_\_\,]_{\text{wd}}$

(3/7: *sap, fet, gluk, *drup, *kop, *rat, *tik*)

$p \to bi$ rules

$\text{p} \to \text{bi} \,/\, \text{dru} \,\_\_\,]_{\text{wd}}$

$\text{p} \to \text{bi} \,/\, \text{ko} \,\_\_\,]_{\text{wd}}$

$\text{p} \to \text{bi} \,/\, \begin{bmatrix} +\text{syl} \\ -\text{low} \\ +\text{round} \end{bmatrix} \,\_\_\,]_{\text{wd}}$

(2/3: *drup, kop, *sap*)

$t \to di$ rules

$\text{t} \to \text{di} \,/\, \text{ra} \,\_\_\,]_{\text{wd}}$

$a \to i$ rules

$\varnothing \to \text{i} \,/\, \text{zup} \,\_\_\,]_{\text{wd}}$

$\varnothing \to \text{i} \,/\, \text{krom} \,\_\_\,]_{\text{wd}}$

$\varnothing \to \text{i} \,/\, \begin{bmatrix} -\text{syllabic} \\ +\text{labial} \end{bmatrix} \,\_\_\,]_{\text{wd}}$

(2/2: *zupa, kroma*)

$k \to gi$ rules

$\text{k} \to \text{gi} \,/\, \text{ti} \,\_\_\,]_{\text{wd}}$

Learning: PL $\rightarrow$ SG direction

- *i* $\rightarrow \varnothing$ (*sapi* $\sim$ *sap*) vs. *i* $\rightarrow$ *a* (*zupi* $\sim$ *zupa*): obstruents prefer former, just one exception
- A phonological rule of final devoicing is also helpful (see Albright & Hayes, 2002)
- On the whole, much less ambiguity in this direction

Choosing a source of UR

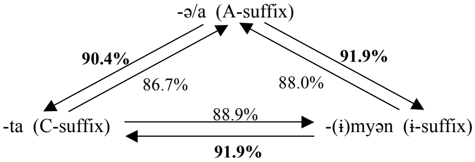
- Goal: find the most predictive form in the paradigm to serve as the input for rules
- Strategy: "wug test" each candidate grammar on the known pairs
  - Given SG. forms *sap*, *rat*, *drup*, *zupa*, etc. use grammar to generate PL., and vice versa
  - For each form, check whether the grammar was right ($\rightarrow$ ACCURACY)
  - Also potentially relevant: how many competitors, difference in reliability
- Form that yields the most accurate outputs is overall most predictive
  - Here, plural more predictive: no need to 'undo' final devoicing, and singular *-a* is predictable for *kroma* (non-obstruent)

Some points to note

- Under condition B'' (Single Surface Base restriction), may not be any perfectly predictive choice
  - Neutralizations in different contexts
  - Numerical competition
- Properties in the base are taken for granted (not predicted), properties in other forms are predicted
  - Source of asymmetries in generalization/innovation

Assessment

- Correctly predicts direction of generalization in various cases
  - Spanish verbs (Albright, 2003)
  - Latin nouns (Albright, 2005)
  - Korean verbs (Albright & Kang, 2009)



  - Yiddish verbs (Albright, 2010)

Condition B vs B″

- Many cases appear to show generalization from unaffixed ('isolation') form, in spite of seemingly rather severe neutralizations
  - Dutch, Turkish, Korean final laryngeals
  - Yidiɲ, Māori apocope
- Albright (2008): a role for token frequency
  - Although predictability doesn't explicitly take token frequency into account, high frequency forms provide data about more lexical items
  - Confidence limits reward rules based on more items
  - Result: easier for grammars using frequent forms as input to achieve higher confidence
- Application: Korean nouns, Dutch nouns, etc.

Beyond Condition B$''$

Wang and Hayes (submitted) Learning underlying representations: The role of abstractness

- Attempt to search larger hypothesis space of UR's
- Parameter of the model: up to which K&K level?
- Model hypothesizes set of candidate UR's, attempts to learn probability distribution over them + weighting of constraints that can generate the observed data

Candidate UR's

- Model receives sets of morphologically related forms, much like MGL

Stem	Plural	Gloss
bet	beda	'cat'
mot	mota	'dog'
lop	loba	'turtle'
pap	papa	'soup'
mik	miga	'plane'
bek	beka	'beer'
es	esa	'wine'
nur	nura	'light'
to	toa	'toe'

- Morpheme segmentation: minimize differences between morphemes
- Similar idea as MGL segmentation, though different

UR hypotheses

- KK level B: just the ‘isolation’ allomorphy
  - /bet/, /mot/, /lop/, /pap/, …
- KK level B': could check for most widely-occurring form (Wang and Hayes don’t dwell on this)
- KK level B'': not implemented
- KK level C: all allomorphs
  - {[bet], [bed]}, {[mot]}, {[lop], [lob]}, {[pap]}, …
- KK level D: recombine segments (here, no new candidate URs)

UR hypotheses

- KK level E: recombination

(28) *Forming the set of candidate URs for Seediq 'hold'*

a. Optimal string alignment

p	**e**	m	**u**	x	isolation allomorph
p	**u**	m	**e**	x	single-suffix allomorph

b. Free combination of alternatives to create UR candidates

/p	e	m	u	x/	
/p	**e**	**m**	**e**	**x/**	(emerges as correct under further learning)
/p	u	m	u	x/	
/p	u	m	e	x/	

- And beyond? KK level "Z"
  - Add candidate UR's that alter segments never seen to alternate, epenthesize, delete, etc.

Selecting URs from among the candidates

- Challenge: simultaneously learn lexicon and grammar
- Lexicon: probability distribution over potential URs
  - Initially flat (all potential URs equally probably)
- Grammar: weighted constraints
  - Constraints given in advance
  - Gen creates candidates from potential URs
  - Eval: violations determine winners under current weights
- Adjusting UR probabilities and constraint weights: Expectation Maximization
  - Expectation: given the current grammar, which URs are doing the best job of predicting the given SRs?
  - Maximization: update constraint weights and re-estimate UR probabilities

Results

- Wang and Hayes show that this model is able to find lexicon/grammar pairings that correctly capture the data with intuitively correct grammars and sufficiently abstract URs
  - Mostly, KK-C (pick an allomorph) and KK-D (recombination)
- This models A answers to problem sets, not necessarily what humans do
- Not considered: conditions under which the model fails, favoring UR candidates corresponding to lower levels?
  - Also, KK-B$''$ not actually implemented

Finding predictability, more broadly

- Wang and Hayes address one form of abstraction: strings that don't occur on the surface
- Not addressed
  - Changing feature values
  - Removing feature values
- Changing: Bakwiri nasal [kõmbà] $\Rightarrow$ oral /kombà/
- Removing: no final voicing specification at all?
- These moves were central to K&K's formulation: "unpredictable properties"

References

Albright, Adam. 2003. A quantitative study of Spanish paradigm gaps. *Wccfl 22 proceedings*, ed. by G. Garding and M. Tsujimura, 1–14. Somerville, MA: Cascadilla Press.

Albright, Adam. 2005. The morphological basis of paradigm leveling. *Paradigms in phonological theory*, ed. by Laura J. Downing, Tracy Alan Hall, and Renate Raffelsiefen, 17–43. Oxford University Press.

Albright, Adam. 2008. Explaining universal tendencies and language particulars in analogical change. *Language universals and language change*, ed. by Jeff Good, 144–181. Oxford University Press.

Albright, Adam. 2010. Base-driven leveling in Yiddish verb paradigms. *Natural Language and Linguistic Theory* 28.475–537.

References (*cont.*)

ALBRIGHT, ADAM, and BRUCE HAYES. 2002. Modeling English past tense intuitions with minimal generalization. *Sigphon 6: Proceedings of the sixth meeting of the acl special interest group in computational phonology*, 58–69. ACL.

ALBRIGHT, ADAM, and YOONJUNG KANG. 2009. Predicting innovative alternations in Korean verb paradigms. *Current issues in unity and diversity of languages: Collection of the papers selected from the cil 18, held at korean university in seoul*, 893–914. Linguistic Society of Korea.