Learning Composite Representations

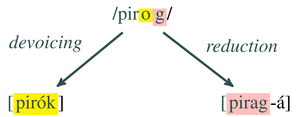
Dustin Bowers (bowersd@ucla.edu)
UCLA Linguistics

What is a Composite UR?

- A composite UR contains features drawn from multiple allomorphs.
- Two allomorphs of morpheme for 'pie' in Russian: *pirók*, *pirag*

	sg.	pl.
nominative	pirók	pirag-í
accusative	pirók	pirag-í
genitive	pirag-á	pirag-óf
dative	pirag-ú	pirag-ám
instrumental	pirag-óm	pirag-ámi
locative	pirag-é	pirag-áx

- UR must be composite to capture:
  - stressed vowel quality (pirók, *pirák)
  - consonant voicing (pirag-á, *pirak-á)



Problem for the Single Surface Base Hypothesis

- Albright (2002; 2010) argues one single slot in paradigm is basis for derivation of all others.
  - Therefore composite UR systems should be impossible.
  - Predicts systems like Russian eventually collapse.

→ The above aspect of Russian has been stable for ∼ 700 years.

- **Goal: learn composite URs while capturing facts that support the Single Surface Base Hypothesis.**

Tesar's Learning Model

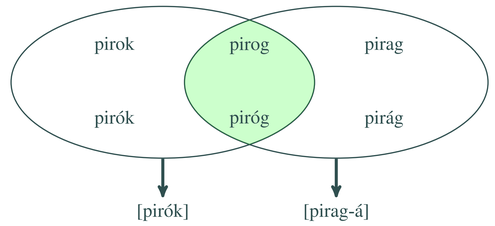
- Tesar's (2013) model rapidly finds composite URs.
  1. Phonotactic ranking.
  2. Set intersection of possible URs for every allomorph.
     - UR → SR may only add to phonotactic ranking.
- For instance, Russian phonotactic ranking produces:

	vrat͡ʃ-á	*VTV	*D#	ID-VOI	*o	*á	ID-LO
a. ☞	vrat͡ʃ-á						
b.	vrad͡ʒ-á	L		W			
c.	vrat͡ʃ-ó					L	W

ID-VOI ≫ *VTV (no inter-V voicing)

ID-LOW ≫ *á (no raising)

- The intersection of possible URs for *pirók*, *pirag* is not empty:



- Intersection URs → either SR, only require further MARK ≫ FAITH:
  - *D# ≫ ID-VOI ≫ *VTV (devoicing, no inter-V voicing)
  - *o ≫ ID-LOW ≫ *á (reduction, no raising)
- **If alternations and phonotactics are consistent, this model learns composite URs.**

What Supports Single Surface Base-ism?

- If paradigm is derived from a single surface form, it might be spread to other slots of the paradigm in language change.
- For example, Middle High German innovated schwa apocope (King 1976, Albright 2008)...
  - Setting up an opaque system ripe for change.

'praise'	'praise-nom.pl'	
/lob/	/lob-ə/	UR
lop	—	Devoicing
—	lo:bə	Open σ Lengthening
—	lo:b	Schwa Apocope
[lop]	[lo:b]	

- Yiddish nouns maintained plural, lost singular allomorphs:

MHG		Pre-Yiddish		Yiddish	
lop	>	*lo:b	>	lɔɪb	'praise'
lo:b-(ə)	>	*lo:b	>	lɔɪb-ən	'praise-pl'

- This is predicted by Single Surface Basism because learners privileged one allomorph at the expense of the other.
  - The plural allomorph generally carried more contrast and thus is predicted to have been selected as base of the paradigm.

Not an Argument against Composite URs

- A composite UR would not have enabled the alternation to be maintained.
- Opacity made formerly allophonic vowel length in closed syllables, obstruent voicing contrastive.
- **The alternations are not consistent with the phonotactics.**

Applying Tesar's Learning Model to Yiddish: No UR

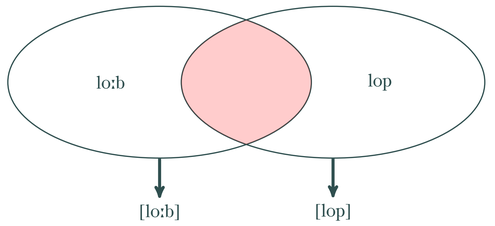
- Phonotactic ranking after apocope:

	lo:b	ID-VOI	*D#	ID-LONG	*V:C(C)]$_\sigma$
a. ☞	lo:b				
b.	lo:p	W	L		
c.	lob			W	L

Id-long ≫ *V:C(C)]$_\sigma$

Id-voi ≫ *D#

- The phonotactic ranking is inconsistent with *lo:b* - *lop* alternation.
- No UR can map to both allomorphs.
  - Possible URs for each allomorph have empty intersection.



Modify Albright's Model: Base as Backup

- If an empty intersection occurs, fall back to URs for a single allomorph.
- Which allomorph to fall back on determined as in Albright's work.
- For Yiddish, this was the plural (see Albright 2008).
- **Result: /lo:b/ is only available UR.**
- Applying grammar to /lo:b/ creates an acquisition error in the singular:
  - Child [lo:b] in place of parental [lop]

	lo:b	ID-VOI	*D#	ID-LONG	*V:C(C)]$_\sigma$
a. ☹	lop	*(!)		*(!)	
b. ☞	lo:b		*		*

- This matches the observed historical change.

Why did it work?

- The opacity of system made former allophones contrastive.
- Tesar's model is too fragile to find the "right" analysis.
- Long vowels and voicing in plural (base) projected into UR.
- Voicing and length spread to singular by faithfulness.

Similar Cases

- Other languages with leveling, and Tesarian UR learning yields an empty set:
  - Odawa and rhythmic syncope systems (Bowers in prep)
  - Latin *honor* analogy.
  - Serbo-croatian (leveling of l∼o alternations)
- Evidently stable languages with composite URs:
  - Palauan
  - Tonkawa
  - Pima
  - New Odawa

Conclusions

- The problem: how to reconcile evidence for single surface basism with need for composite URs.
- Both Tesar and Albright are correct.
  - Reconstrue the single surface base hypothesis as a default backup strategy.
  - Make constructive use of the fragility of Tesar's method.

→ **General prediction: Composite URs are only possible when a complete phonological analysis is available**

  - See also Bermúdez-Otero (2014)

References

Albright, A. (2002). *The Identification of Bases in Morphological Paradigms*. Ph. D. thesis, University of California, Los Angeles.

Albright, A. (2008). Inflectional paradigms have bases too: Evidence from Yiddish. In A. Bachrach and A. Nevins (Eds.), *The Bases of Inflectional Identity*. Oxford University Press.

Albright, A. (2010). Base-driven leveling in Yiddish verb paradigms. *Natural Language and Linguistic Theory 28*, 475–537.

Bermúdez-Otero, R. (2014). French adjectival liaison: Evidence for underlying representations. Handout Given at Oxford University.

King, R. (1976). *The History of Final Devoicing in Yiddish*. Indiana University Linguistics Club.

Tesar, B. (2013). *Output-Driven Phonology: Theory and Learning*. Cambridge University Press.

Acknowledgements

Thanks to Bruce Hayes, Kie Zuraw, Ed Stabler, audiences at UCLA and the Annual Meeting of Phonology (2014).