Michael Becker (becker@fas.harvard.edu) February 13–14, 2010
Computational Modelling of Sound Pattern Acquisition University of Alberta

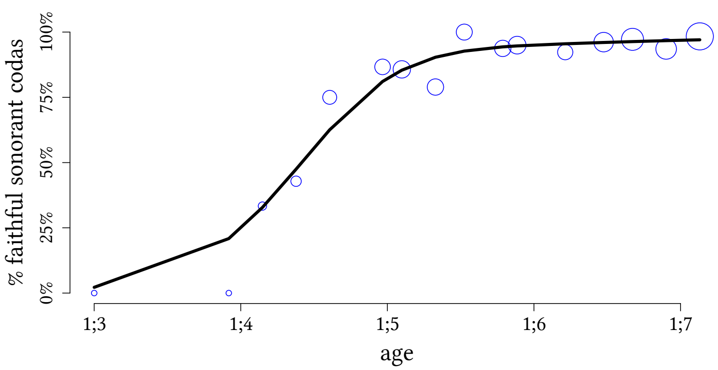
Target selection in error-based learning[*]

- SR (AKA ˈʃaχaʁ̥, Adam & Bat-El 2007a,b, 2009, 2010) shows a typical acquisition pattern of first *avoiding* marked structures, then *repairing* them, then producing them faithfully.
- I offer an analysis (w/ computational implementation) of SR's acquisition path in terms of error-based learning, modeling the avoidance of marked structure as selection of the null parse.
- The model predicts a realistic acquisition path, given a persistent M > F > MPARSE bias, if MPARSE is relativized to Markedness constraints.

1 What is target selection?

SR starts by deleting sonorant codas, then fairly quickly produces them faithfully:

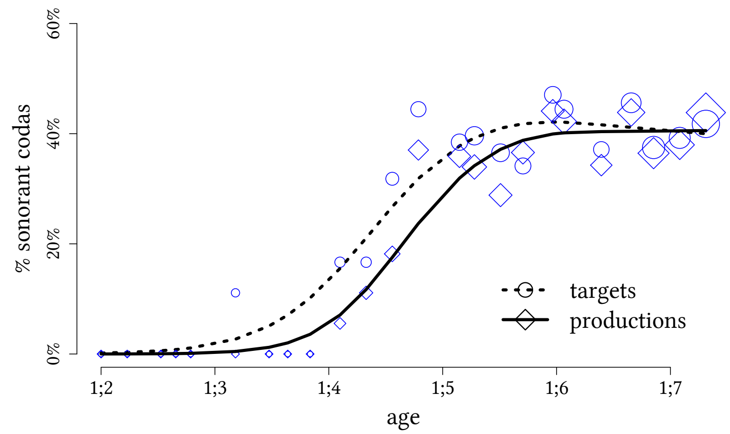
(1) SR faithfulness to sonorant codas



[*]Thanks to Outi Bat-El for sharing her data and her insights. Thanks also to Adam Albright, Hannes Fellner, Wendell Kimper, Joe Pater, John McCarthy, and Matt Wolf for feedback and valuable comments. Any remaining errors are due to partial or imperfect knowledge and flawed reasoning.

But if we look at the proportion of sonorant codas out of all codas, and start a month earlier, we see that SR usually avoids words with sonorant codas, and only rarely deletes their codas:[1]

(2) SR's development of sonorant codas



(3) SR carves out an increasingly representative subset of the adult language for his productions, gradually approaching a language with ∼40% items with sonorant codas.

(4) Kager et al. (2004, p. 11) quote Kiparsky & Menn (1977, pp. 56–58):

> Different children exclude **definable classes of output by different means** ... the various rules of child phonology (substitutions, deletions, etc.) as well as **selective avoidance of some adult words**, are devices the child finds for dealing with those difficulties.

(5) In OT terms, a single Markedness constraint against sonorant codas causes both deletion of these codas and avoidance of words that have such codas in their adult form.

[1]Adam & Bat-El (2007a,b, 2009, 2010) discuss SR's target avoidance in the context of prosody. Target avoidance, or selection, is mentioned as a feature of language acquisition in, among others, Ferguson & Farwell (1975); Kiparsky & Menn (1977); Schwartz & Leonard (1982); Stoel-Gammon & Cooper (1984); Fikkert (1994); Grijzenhout & Joppen-Hellwig (2002); Menn (2004); Goad & Rose (2004).

2 Target selection: Syllable margins

2.1 The data

SR usually repairs sonorant codas by deletion (16/30 unfaithful cases)[2]

(6)

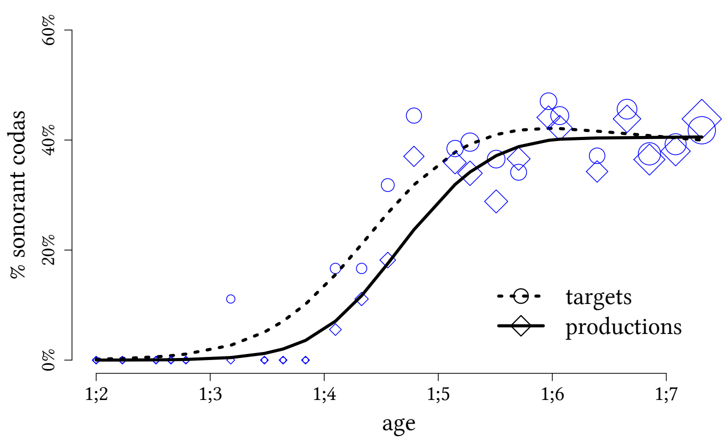
Adult target	SR's production		
ˈpil	ˈpi	1;04.10	'elephant'
na.ˈmeʁ̞	na.ˈme	1;05.08	'tiger'
ˈma.im	ˈma.i	1;05.15	'water'

(7) *Unfaithful* productions of sonorant codas start at 1;4.03 (+one at 1;3.05). *Faithful* productions of sonorant codas start at 1;4.03.

(8) In other words, avoidance is SR's preferred approach, with a short unfaithful lag between avoidance and adult forms.

(9) Avoidance (up to ∼1;5) → deletion (∼1;4–1;5) → faithfulness (∼1;5—)

(10) SR's development of sonorant codas (repeated from (2))



[2]SR opts for metathesis in 12/30 cases when it repairs an onsetless unstressed syllable, as in [ˈja.li] or [ˈla.li] for adult [ˈla.il] 'Lyle'.

SR usually repairs complex onsets by deletion (22/27 unfaithful cases)

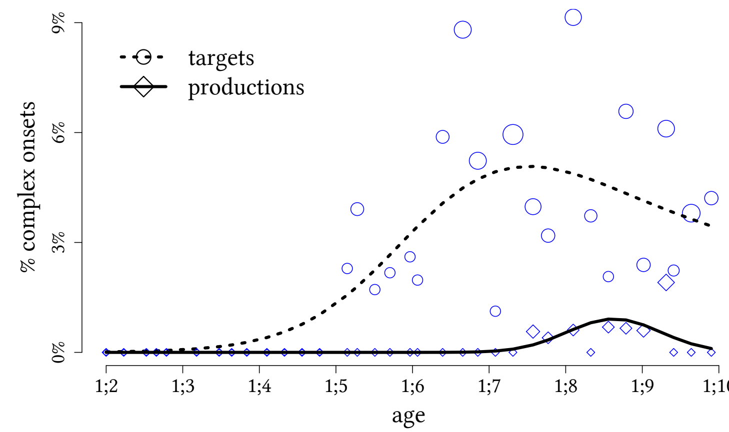
(11)

Adult target	SR's production		
ˈpkak	ˈpak	1;05.29	'cap'
ˈgli.da	ˈgi.da	1;06.12	'ice-cream'
dvo.ˈʁ̞a	do.ˈʁ̞a	1;06.26	'bee'

(12) First *unfaithful* production of a complex coda at 1;5.04. First *faithful* production at 1;7.17 – more than two months later.

(13) Avoidance (up to ∼1;5) → deletion (∼1;5–1;7) → faithfulness (∼1;7—)

(14) SR's development of complex onsets



2.2 Analysis: Error-based learning of target avoidance

Avoidance = the null parse is the winner.

(15) The null parse violates MParse, but no Markedness or Faithfulness constraints (Prince & Smolensky 1993/2004; McCarthy & Wolf 2010).

(16) Usually used to model paradigm gaps (e.g. Rice 2006, Bat-El 2010, but see Albright 2010)

(17) Extended to phonotactic gaps in Prince & Smolensky (1993/2004); Smith (2009)

Paired with error-based learning (Prince & Tesar 2004; Hayes 2004 et seq.):

(18) $\mathcal{H}_0$: M ≫ F ≫ MParse
In the absence of evidence, MParse is at the bottom, so everything is avoided. The child is silent, as is indeed observed.

(19)

	M	F	MParse
a. adult form ≻ ⊙	L		W

$\mathcal{H}_1$: MParse ≫ M ≫ F
The adult form and ⊙ never violate Faithfulness, so pairing them necessarily allows ranking MParse over M.

(20)

	M	F	MParse
a. adult form ≻ ⊙	L		W
b. adult form ≻ simplified form	L	W	

(21) $\mathcal{H}_2$: F ≫ MParse ≫ M = adult grammar

2.3 Analysis: application to syllable margins

Acquisition of sonorant codas:

(22)

/na.ˈmeʁ̥/	*SonCoda	Max-C	MParse
a. na.ˈmeʁ̥	*!		
b. na.ˈme		*!	
c. ☞ ⊙			*

(23)

Avoidance	$\mathcal{H}_0$ (up to ∼1;5):	*SonCoda ≫ Max-C ≫ MParse
Deletion	$\mathcal{H}_1$ (∼1;4–1;5):	MParse ≫ *SonCoda ≫ Max-C
Adult	$\mathcal{H}_2$ (∼1;5—):	Max-C ≫ MParse ≫ *SonCoda

Acquisition of complex onsets:

(24)

Avoidance	$\mathcal{H}_0$ (up to ∼1;5):	*ComplexOnset ≫ Max-C ≫ MParse
Deletion	$\mathcal{H}_1$ (∼1;5–1;9):	MParse ≫ *ComplexOnset ≫ Max-C
Adult	$\mathcal{H}_2$ (∼1;9—):	Max-C ≫ MParse ≫ *ComplexOnset

But both acquisition paths depend on the ranking of Max-C and MParse:

(25)

$\mathcal{H}_0$ (up to ∼1;4):	*ComplexOnset, *SonCoda ≫ Max-C ≫ MParse
$\mathcal{H}_1$ (∼1;4–1;5):	*ComplexOnset ≫ MParse ≫ *SonCoda ≫ Max-C
$\mathcal{H}_2$ (∼1;4–1;5):	*ComplexOnset ≫ Max-C ≫ MParse ≫ *SonCoda
$\mathcal{H}_3$ (∼1;5–1;7):	MParse ≫ *ComplexOnset ≫ Max-C ≫ *SonCoda
$\mathcal{H}_4$ (∼1;7—):	Max-C ≫ MParse ≫ *ComplexOnset ≫ *SonCoda

(26)

$\mathcal{H}_0$ (up to ∼1;4):	Complex onsets: avoided	Sonorant codas: avoided
$\mathcal{H}_1$ (∼1;4–1;5):	Complex onsets: ***simplified***	Sonorant codas: simplified
$\mathcal{H}_2$ (∼1;4–1;5):	Complex onsets: avoided	Sonorant codas: faithful
$\mathcal{H}_3$ (∼1;5–1;7):	Complex onsets: simplified	Sonorant codas: faithful
$\mathcal{H}_4$ (∼1;7—):	Complex onsets: faithful	Sonorant codas: faithful

When MParse outranks Max-C, avoidance is no longer acceptable for either sonorant codas or complex onsets. It doesn't matter that different markedness constraints are involved.

(27) We don't observe these simplified complex onsets during the relevant period (∼1;4–1;5).

(28) Due to the short relevant period (<1 month) and the general paucity of complex onsets, the number of simplified complex onsets is expected to be ∼2, which is not significantly different from zero (Fisher's Exact test, $p > .1$).

(29) And yet, there is reason to think that the prediction in (25-25) in wrong.

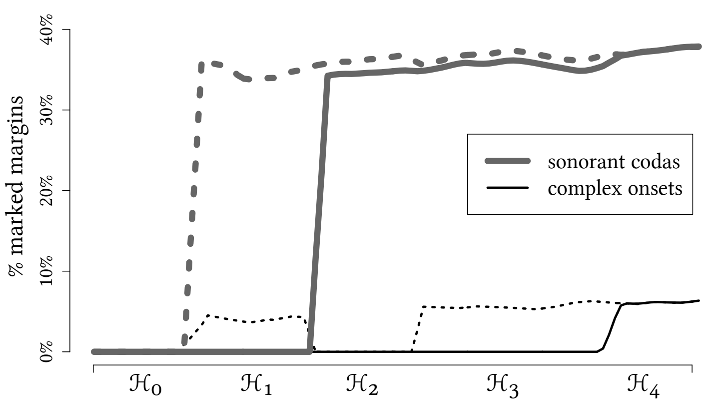
2.4 Implementation

The learner (a Perl script, available at http://becker.phonologist.org/shaxar/, based on Becker & Tessier 2010; Becker 2005):

- CON (*SonCoda, *ComplexOnset, Max, MParse) à la Becker (2005)
- GEN makes candidates for each input (faithful, ⊙, simplified onset/coda)
- EVAL chooses an output
- Error are stored in the Cache (Tessier 2007, 2008, 2009; Becker & Tessier 2010)
- Errors from the Cache are selectively moved to the Support
- Biased Constraint Demotion Algorithm (à la Prince & Tesar 2004; Hayes 2004) creates a new grammar, persistently favoring M > F > MParse.

The learner runs a list of SR's targets through its current grammar, one by one. When enough errors accumulate, the Support is updated and a new grammar is learned.

(30) Model predictions: Productions and attempts of marked syllable margins



(31) The language's more abundant marked structures will cause errors to pile up quickly, causing the model to acquire those structures first.

(32) Sonorant codas are about 3–6 times more common than complex onsets in Hebrew, yet SR acquires complex onsets very shortly after sonorant codas, i.e. the frequency effect is too big. I enable the model to mimic SR's behavior by allowing learning thresholds to differ by Markedness constraint.

3 Targer selection with MParse(M)

3.1 The problem

Modeling target selection as an MParse effect, paired with a persistent M > F > MParse bias, predicts that avoidance will be turned on and off all the time:

(33) $\mathcal{H}_0$: $M_1, M_2, M_3, \ldots \gg F_1, F_2, F_3, \ldots \gg$ MParse
Initial state: everything avoided.

(34)

	M_1	M_2	M_3	F_1	F_2	F_3	MParse
a. adult ≻ ⊙	L						W

$\mathcal{H}_1$: $M_2, M_3, \ldots \gg$ MParse $\gg M_1 \gg F_1, F_2, F_3, \ldots$
Nothing avoided; everything simplified.

(35)

	M_1	M_2	M_3	F_1	F_2	F_3	MParse
a. adult ≻ ⊙	L						W
b. adult ≻ simplified	L			W			

$\mathcal{H}_2$: $M_2, M_3, \ldots \gg F_1 \gg$ MParse $\gg M_1 \gg F_2, F_3, \ldots$
M_1-violating forms surface faithfully; everything else avoided.

This seems like a rather unappealing model of acquisition.

3.2 The solution: Markedness-based avoidance

Avoidance is a speaker's mechanism for dealing with a Markedness violation. Instead of promoting MParse as in §3.1, promote-markedness specific instances of MParse.[3]

(36) $\mathcal{H}_0$: $M_1, M_2, M_3, \ldots \gg F_1, F_2, F_3, \ldots \gg$ MParse
Initial state: everything avoided.

(37)

	M_1	M_2	M_3	F_1	F_2	F_3	MParse
a. adult ≻ ⊙	L						W

$\mathcal{H}_1$: $M_2, M_3, \ldots \gg$ MParse(M_1) $\gg M_1 \gg F_1, F_2, F_3, \ldots \gg$ MParse
M_1-violating adult forms simplified; everything else avoided.

(38)

	M_1	M_2	M_3	F_1	F_2	F_3	MP(M_1)	MP
a. adult ≻ ⊙	L						W	W
b. adult ≻ simplified	L			W				

$\mathcal{H}_2$: $M_2, M_3, \ldots \gg F_1 \gg$ MParse(M_1) $\gg M_1 \gg F_2, F_3, \ldots \gg$ MParse
M_1-violating forms no longer avoided or simplified; everything else avoided.

MParse(M_1) is defined relative to the fully faithful candidate (McCarthy 2003, 2007), which in this acquisition stage is conveniently identical to the adult form (as in, e.g. Hayes 2004).

(39) MParse(M_1): Assign a violation mark to ⊙ if the FFC violates M_1

And more generally:

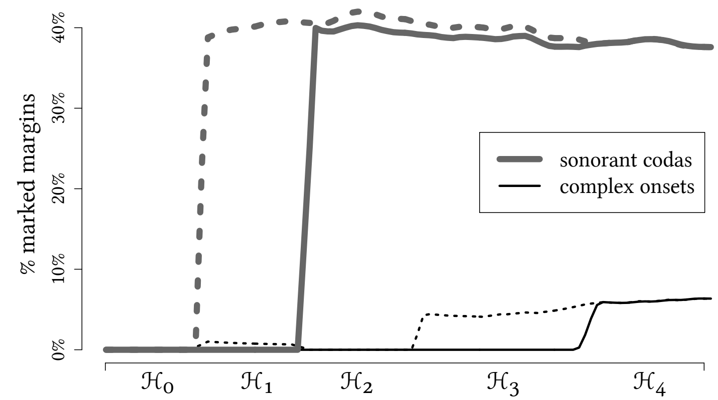
(40) MParse($M_m \ldots M_n$): Assign a violation mark to ⊙ if the FFC violates all of the Markedness constraints in $M_m \ldots M_n$

[3] More specific versions of MParse were proposed in Rice (2006).

3.3 Implementation, take 2

The model is based on the one in §2.4, but now it installs Markedness-specific versions of MParse. The original MParse stays at the bottom of the hierarchy.

(41) Model predictions: Productions and attempts of marked syllable margins



(42) Still a few stray complex onsets at $\mathcal{H}_1$, but not nearly as many as before.

(43) $\mathcal{H}_1$: *ComplexOnset $\gg$ MParse(*SonC) $\gg$ *SonCoda $\gg$ Max-C $\gg$ MParse

Simplifying a complex onset or a sonorant coda is still done by deletion, so the same Max-C is violated in both cases. What happens in $\mathcal{H}_1$ when the input has both?

(44)

/pʁ̥a.ˈχim/	*Complex Onset	MParse (*SonC)	*SonCoda	Max-C	MParse
a. pʁ̥a.ˈχim	*!		*		
b. pa.ˈχim			*!	*	
c. ☞ pa.ˈχi				**	
d. ⊙		*!			*

Now two marked configurations that involve the same faithfulness constraint are predicted to interact, but *only in the words that have both configurations.*

4 Conclusions

- I analyzed SR's (Adam & Bat-El 2007a,b, 2009, 2010) acquisition of marked syllable margins, showing that he first *avoids* words that have them, then he *deletes* the offending consonants, then produces them faithfully.
- I offered an analysis (w/ computational implementation) of SR's acquisition path in terms of error-based learning, modeling the avoidance of marked structure as selection of $\odot$. The model includes a persistent M > F > MParse bias and relativization of MParse to Markedness constraints.
- Choosing the null parse emerges as a convenient output, but more importantly, an informative one: It tells the learner which Markedness constraint to demote.
- The M > F > MParse calculus is rather micro-manage-y. To be replaced with something more elegant?...

References

Adam, Galit & Outi Bat-El (2007a). Segmental effects on syllable selection: Evidence from Hebrew. Handout from Generative Approaches to Language Acquisition, Barcelona.

Adam, Galit & Outi Bat-El (2007b). The trochaic bias is universal: Evidence from Hebrew. Handout from Generative Approaches to Language Acquisition, Barcelona.

Adam, Galit & Outi Bat-El (2009). When do universal preferences emerge in language development? The acquisition of Hebrew stress. *Brill's Annual of Afroasiatic Languages and Linguistics* **1**. 255–282.

Adam, Galit & Outi Bat-El (2010). The trochaic bias in the acquisition of Hebrew. Ms. Tel Aviv University.

Albright, Adam (2010). Lexical and morphological conditioning of paradigm gaps. In *Modeling Ungrammaticality in Optimality Theory*, London: Equinox.

Bat-El, Outi (2010). A gap in the feminine paradigm of Hebrew: A consequence of identity avoidance in the suffix domain. In Curt Rice (ed.) *When Nothing Wins: Modelling Ungrammaticality in OT*, London: Equinox.

Becker, Michael (2005). CCamelOT - an implementation of OT-CC's gen and eval in Perl. Software, UMass Amherst, http://people.umass.edu/mbe/CCamelOT/.

Becker, Michael & Anne-Michelle Tessier (2010). Trajectories of faithfulness in child-specific phonology. Talk presented at the 84th Annual Meeting of the LSA, Baltimore, MD.

Ferguson, Charles A. & Carol Farwell (1975). Words and sounds in early language acquisition. *Language* **51**. 419–439.

Fikkert, Paula (1994). *On the Acquisition of Prosodic Structure.* Ph.D. dissertation, Leiden University.

Goad, Heather & Yvan Rose (2004). Acquisition of left-edge clusters in west germanic. In René Kager, Joe Pater & Wim Zonneveld (eds.) *Constraints in Phonological Acquisition*, Cambridge University Press. 109–157.

Grijzenhout, Janet & Sandra Joppen-Hellwig (2002). The lack of onsets in german child phonology. In I. Lasser (ed.) *The Process of Language Acquisition*, Frankfurt/Berlin: Peter Lang Verlag. 319–339.

Hayes, Bruce (2004). Phonological acquisition in optimality theory: the early stages. In René Kager, Joe Pater & Wim Zonneveld (eds.) *Constraints in Phonological Acquisition*, Cambridge: Cambridge University Press. 158–203.

Kager, René, Joe Pater & Wim Zonneveld (2004). *Constraints in Phonological Acquisition.* Cambridge University Press.

Kiparsky, Paul & Lise Menn (1977). On the acquisition of phonology. In *Language Learning and Thought*, New York: Academic Press. 47–78.

McCarthy, John J. (2003). Comparative markedness. *Theoretical Linguistics* **29**. 1–51.

McCarthy, John J. (2007). *Hidden Generalizations: Phonological Opacity in Optimality Theory*. London: Equinox.

McCarthy, John J. & Matthew Wolf (2010). Less than zero: Correspondence and the null output. In Curt Rice (ed.) *Modeling Ungrammaticality in Optimality Theory*, London: Equinox.

Menn, Lise (2004). Saving the baby: making sure that old data survive new theories. In René Kager, Joe Pater & Wim Zonneveld (eds.) *Constraints in Phonological Acquisition*, Cambridge University Press. 54–72.

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

Prince, Alan & Bruce Tesar (2004). Learning phonotactic distributions. In René Kager, Wim Zonneveld & Joe Pater (eds.) *Fixing Priorities: Constraints in Phonological Acquisition*, Cambridge University Press. 245–291.

Rice, Curt (2006). Nothing is a phonological fact: Gaps and repairs at the phonology-morphology interface. Talk given at NELS 36.

Schwartz, Richard G. & L. Leonard (1982). Do children pick and choose? an examination of phonological selection and avoidance in early lexical acquisition. *Journal of Child Language* **9**. 319–336.

Smith, Brian (2009). Cumulative constraint interaction and the null parse. Talk given at HUMDRUM '09 and MUMM '09.

Stoel-Gammon, Carol & Judith A. Cooper (1984). Patterns of early lexical and phonological development. *Journal of Child Language* **11**. 247–271.

Tessier, Anne-Michelle (2007). *Biases and stages in phonological acquisition.* Ph.D. dissertation, University of Massachusetts, Amherst, Amherst, MA.

Tessier, Anne-Michelle (2008). Children's exceptional words: Lexical challenges from the time-course of L1 phonological acquisition. Talk given at the University of Western Ontario.

Tessier, Anne-Michelle (2009). Frequency of violation and constraint-based phonological learning. *Lingua* **119**. 6–38.