Distance-based decay in long-distance phonological processes

Jesse Zymet
University of California, Los Angeles

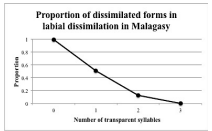
1. Introduction

- Long-distance phonological processes display *distance-based decay*: rate of application decreases as transparent distance between the trigger and target increases.
- This is challenging for approaches to assimilation and dissimilation in which distance does not play a role.
- My account uses the maximum entropy framework (Smolensky and Legendre 2006, et seq.) and a decay function to account for the effect in three processes across four languages.
- The account is implementable in ABC (Hansson 2001, et seq.) with weighted constraints and probabilistic output.

2. Empirical background

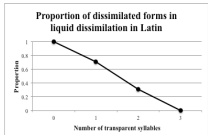
Distance-based decay can be observed in a variety of different long-distance phonological processes:

Vowel dissimilation in Malagasy: /uσ^n+u/ → [uσ^n-i]			
Transparent syllables (σ)	Faithful forms	Dissimilated forms	% of dissim'd forms
$n = 0$	4	989	99.6%
$n = 1$	196	201	50.6%
$n = 2$	28	4	12.5%
$n = 3$	4	0	0%



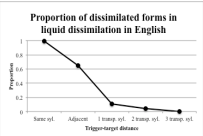
(Local 1) /ba.bu.+u/ → [ba.bu.-**i**]
(Local 2) /tu.r+u/ → [tu.r-**i**]
(Nonlocal 1) /ru.va.+u/ → [ru.va.-u]
(Nonlocal 2) /un.dan.+u/ → [un.dan.-**i**]

Liquid dissimilation in Latin: /lσ^n+a:lis/ → [lσ^n-a:**ris**]			
Transparent syllables (σ)	Faithful forms	Dissimilated forms	% of dissim'd forms
$n = 0$	0	131	100%
$n = 1$	20	49	71.0%
$n = 2$	29	13	31.0%
$n = 3$	4	0	0.0%



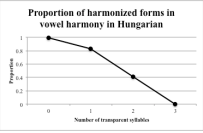
(Local 1) /so.l+a:.lis/ → [so.l-**a:.ris**]
(Local 2) /mu.l+a:.lis/ → [mu.l-**a:.ris**]
(Nonlocal 1) /pa.le.+a:.lis/ → [pa.le.-a:.lis]
(Nonlocal 2) /a.le.+a:.lis/ → [a.le.-**a:.ris**]

Liquid dissimilation in English: /lσ^n+əl/ → [lσ^n-**əɹ**]			
Trigger-target distance	Faithful forms	Dissimilated forms	% of dissim'd forms
Same syllable	1	303	99.7%
Adjacent syl.s	60	39	65%
1 transp. syl.	85	10	10.5%
2 transp. syl.s	24	1	4.0%
3 transp. syl.s	4	0	0.0%



(Local 1) /soo.l+əl/ → [soo.l-**əɹ**]
(Local 2) /kan.da.l+əl/ → [kan.da.l-**əɹ**]
(Nonlocal 1) /sə.laj.v+əl/ → [sə.laj.v-əl]
(Nonlocal 2) /kə.ləm.n+əl/ → [kə.ləm.n-**əɹ**]

Vowel harmony in Hungarian: /Bσ^n+nək/ → [Bσ^n-**nək**]			
Transparent syllables (σ)	Faithful forms	Harmonized forms	% of harmn'd forms
$n = 0$	4.32	6284.68	99.9%
$n = 1$	128.04	633.96	83.2%
$n = 2$	60.79	42.35	41.1%
$n = 3$	8	0	0%



(Local 1) /əb.lək.+nək/ → [əb.lək.-**nək**]
(Local 2) /bi.ro:.+nək/ → [bi.ro:.-**nək**]
(Nonlocal 1) /kə.li.be:r.+nək/ → [kə.li.be:r.-nək]
(Nonlocal 2) /bo.ri:.te:k.+nək/ → [bo.ri:.te:k.-**nək**]

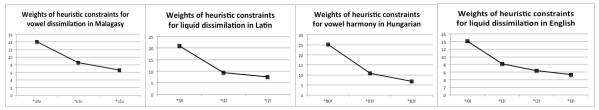
3. Statistical analysis

Likelihood ratio tests over generalized linear models of the data reveal the following:

- The **number of transparent syllables comes out significant** in producing the decay effect for all of the covered languages ($p < 0.001$ for all languages except Latin, for which $p = 0.03$).
- When number of transparent segments was pit against number of transparent syllables, the **number of transparent segments did not come out significant**.
- In Latin, triggers in onset-noninitial position are associated with lower application rates ($p < 0.001$), while labial consonants and velar consonants were not significant (contra Cser 2010).
- In English, triggers in onset-nonintial position or in coda position are associated with lower application rate ($p < 0.001$); velar consonants were associated with lower application rates ($p = 0.02$), while labial consonants were not.
- In Hungarian, height of transparent vowels is significantly related to likelihood of application ($p < 0.001$) as noted by Hayes and Londe 2008.

4. Account rationale

- Positing distance-based markedness constraints is too powerful: learners never acquire small weights for local constraints and large weights for nonlocal ones.
- My account is an extension of Kimper 2011, who proposes a better option: posit one nonlocal markedness constraint whose weight is then scaled with distance.



- Distance-based constraints show that the weight of markedness of cooccurrences decreases with distance in an inverse-exponential fashion (also predicted by Kimper 2011).
- I posit $d(x) = 1/x^k$,where x is syllabic distance (0 if the trigger and target are in the same syllable, 1 if they are adjacent, etc.), and k is a real-valued parameter.
- E.g., in Latin, $d(x)$ scales the weight of markedness as follows:

/…lσ^x+alis/ (e.g., /lapid+a:lis /)	*1…1 $w = 10.97$	IDENT([lat]) $w = 4.53$	Score	Harmony	Predicted probability	Observed probability
[…lσ^x-a:lis]	$1/3^{1.0}$ ≈ 0.33 violations		$10.97 * 0.33$	$e^{-10.97*0.33}$	$\frac{e^{-10.97*0.33}}{(e^{-10.97*0.33} + e^{-4.53*1})} \approx 0.70$	0.69
[…lσ^x-**a:ris**]		1	$4.53 * 1$	$e^{-4.53*1}$	$\frac{e^{-4.53*1}}{(e^{-10.97*0.33} + e^{-4.53*1})} \approx 0.30$	0.31

5. Fitting the model parameters

- Three parameter values must be found: the weight of markedness, the weight of faithfulness, and the decay parameter k.
- I found parameter values that minimize the summed difference between the observed and predicted probabilities.
- As it turns out, **we do not need language-specific values of k**.
- Setting $k = 0.98$ constant across languages and letting the learner determine weights leads to a model with minimal error:
- My current findings suggest that swapping the inverse exponential model for a linear model results in compromised fit.

Vowel dissimilation in Malagasy		Liquid dissimilation in Latin	
w_m :	11.60	w_m :	10.31
w_f :	5.86	w_f :	4.34
k :	**0.98**	k :	**0.98**
Error:	0.01	Error:	0.01
Vowel harmony in Hungarian		**Liquid dissimilation in English**	
w_m :	26.78	w_m :	10.86
w_f :	15.79	w_f :	5.94
k :	**0.98**	k :	**0.98**
Error:	0.78	Error:	0.02

6. Implementation in ABC

- We can scale the weight of correspondence using $d(x)$.
- Imagine we want to account for a consonant harmony process. Suppose that the value of $d(x)$ at two transparent syllables is 0.5.

/C'σ^{2}C/	AGREE[F] $w = 24$	CORR-CC[F] $w = 8$	IDENT([F]) $w = 4$	Score	Harmony	Predicted probability
[C$_1$'σ^{2}C$_1$]	1			$24 * 1 = 24$	e^{-24}	$\frac{e^{-24}}{(e^{-24} + 2e^{-4})} \approx 0.00$
[C$_1$'σ^{2}C$_2$]		0.5		$8 * 0.5 = 4$	e^{-4}	$\frac{e^{-4}}{(e^{-24} + 2e^{-4})} \approx 0.50$
[Cσ^{2}C$_1$]			1	$4 * 1 = 4$	e^{-4}	$\frac{e^{-4}}{(e^{-24} + 2e^{-4})} \approx 0.50$

7. Sources of data

Malagasy:	Beaujardière 1994; www.malagasyword.org
Latin:	The Perseus Digital Library
English:	The Oxford English Dictionary
Hungarian:	Hayes, Zuraw, Siptar, and Londe 2006

8. Selected references

Hansson, G. 2001. Theoretical and typological issues in consonant harmony. Doctoral dissertation, UC Berkeley.

Kimper, W. 2011. Competing triggers: Transparency and opacity in vowel harmony. Doctoral Dissertation, University of Massachusetts Amherst.

Smolensky, P. & Legendre, G. 2006. The harmonic mind: from neural computation to Optimality-theoretic grammar. Cambridge, Massachusetts: MIT Press.