

*Elliott Moreton<sup>1</sup>*  
*University of North Carolina, Chapel Hill*

---

## 1. Introduction

- (1) Main questions:
  - a. How does native-language phonotactics intervene in the perceptual process?
  - b. Given the hypothesis of Optimality-Theoretic phonology, what do those effects tell us? How can we model them?
- (2) Contrast two hypotheses:
  - a. Autonomous perception grammar: Perception of non-lexical stimuli (nonwords) employs a grammar which has no access to the input-output map used in production.
  - b. Analysis by synthesis: Perceptual processes *can* consult the production grammar for legality judgments.  
OT implementation: A constraint FIXPT which penalizes perceptual candidates that are unfaithfully mapped in production.
- (3) Experimental evidence supporting analysis by synthesis: Direct comparison in which perceptual markedness is pitted against production legality.
- (4) Outline of the talk:
  - §2 Autonomous perception grammar
  - §3 Autonomous account for perceptual phonotactic effects
  - §4 "Analysis by synthesis" and FIXPT
  - §5 Experiment
  - §6 Alternative explanations
  - §7 Discussion

---

## 2. Autonomous perception grammar

- (5) A *perception grammar* is a grammar which accepts as input an early perceptual representation (auditory, featural, etc.), and produces as output a prelexical phonological representation—i.e., one which is not restricted to representing existing lexical items.

Terminology follows Boersma (1997, 1999); in particular, distinguishing perception grammar from *recognition grammar*, which handles lexical access. I will be ignoring the lexicon entirely in this talk.

---

<sup>1</sup> I am indebted to many people for discussion of the ideas and facts presented here, including Paul Boersma, Dani Byrd, John Kingston, Yu Li, Jennifer L. Smith, Rachel Walker, and audiences at the 2004 Acoustical Society of America meeting in New York City and at the University of North Carolina at Chapel Hill. Thanks are also due to Chris Wiesen of UNC's Odum Institute for statistical advice. Remaining errors are mine. Correspondence may be addressed to moreton@unc.edu.

- (6) Reasons to think that grammar is used in perception:
- a. Lexical representations (URs) used for both production and comprehension  $\Rightarrow$  need to undo language-particular allophonic and morphophonemic variation in order to match auditory-phonetic representation up with lexical entry (Miller & Chomsky 1963:310–313; Pisoni & Luce 1987).
  - b. Non-native stimuli are perceived as native  $\Rightarrow$  perceptual mechanisms need to represent native-language phonotactic legality (Studdert-Kennedy 1974; Frazier 1987).
- (7) Early proposals assumed that the grammar used in perception was the same as the one used in production, or at most a simple inversion of it:
- "Moreover, [this hypothesis] avoids the implausible assumption that there is one kind of grammar for the talker and another kind for the listener." (Miller & Chomsky 1963:318).
- (8) However, there are empirical arguments that the production and perception grammars may differ in more than just their direction of application:
- a. Acquisition: Child lexical representations can simultaneously differ from adult outputs and from the child's own output. Within-child systematicity  $\Rightarrow$  a rule or constraint component in between the adult output and the child's lexicon, different from the one between the child's lexicon and output (Ingram 1974; Braine 1976; Macken 1980).
  - b. Loanword phonology: Perceptual interpretation of non-native acoustic cues as native categories can occur before *or after* the application of native-language phonological rules, suggesting that it is a grammatical process—and clearly it is not a part of production (Silverman 1992).
  - c. Perception of non-native stimuli Listeners can accurately perceive non-native stimuli which would be repaired in production (Kabak & Idsardi 2003).
- $\Rightarrow$  The grammar used in perception is *similar to*, but not necessarily *identical with*, the one used in production.
- (9) In OT models, this partial overlap is obtained through partial constraint sharing:
- a. Same constraints in both directions: Smolensky 1996 (with markedness applying vacuously in perception); D. Ohala 1996 (ditto); Kenstowicz 2001 (same constraints, different ranking); see also Kirchner 1996.
  - b. Shared faithfulness, different markedness: Lassette & Donegan 1998 (UR-specific markedness); Donegan 2001; Boersma & Hamann 2007 (shared cue constraints)
  - c. Shared markedness, different faithfulness: Pater 1998, 2004.
  - d. Wholly different constraints in both directions: Boersma 1997a, 1998, 1999, 2000, 2001, 2005 (with implicit use of perception grammar in production)
- (10) However, in all of these models the perception grammar is autonomous in the sense that it has no access to the outputs of the production grammar.
- $\Rightarrow$  If a representation is illegal in both perception and production, that is because each grammar, separately, makes it illegal.

---

### 3. Autonomous account of perceptual phonotactics

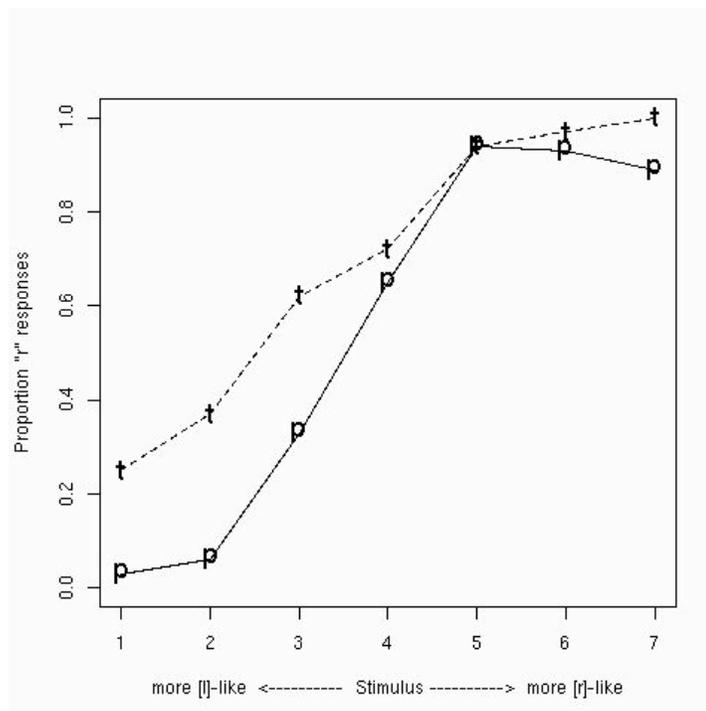
- (11) This section illustrates how autonomous perception grammar can account for experimental data showing perceptual bias...
- a. ... in favor of a legal production output and against an illegal production output, and
  - b. ... towards the "more illegal" of two illegal production outputs.

- (12) Examples will be drawn from the area of L1 phonotactic effects, which are widespread:
- Acoustically *unambiguous* non-native sequences are transcribed as, judged as, or confused with native ones (Polivanov 1931; Brown & Hildum 1956; Messer 1967; Hallé, Segui, Frauenfelder, & Meunier 1998; Brown & Matthews 2001; Kabak & Idsardi 2003; Berent, Steriade, Lennertz, & Vaknin 2006; Berent, Steriade, & Lennertz 2007).
  - Acoustically *ambiguous* phonemes are judged so that they are natively context-appropriate (Massaro & Cohen 1983; Pitt 1998; Dupoux, Kakehi, Hirose, Pallier, & Mehler 1999; Moreton & Amano 1999; Moreton 2002).

### 3.1. Legal parse vs. illegal parse

(13) Massaro & Cohen (1983) experiment:

- Ambiguous segments between [l] and [r]
- Presented to English speakers in contexts [p\_] and [t\_] (among others). Initial [tl] is illegal in English; [tr pr pl] are legal.
- More "r" responses in [p\_] than [t\_] ⇒ English phonotactics affects perception. (Figure replotted from Massaro & Cohen 1983, Figure 2, right-hand panel).



(14) Modelling in terms of an autonomous perception grammar:

- Follow Pater (1998, 2004) in assuming
  - [phonetic] and /phonological/ representations are in discrete featural terms.
  - The same markedness constraints and ranking apply in both directions.
  - Separate faithfulness constraints for production and perception.
 ⇒ The perception grammar has access to a lot of information about the production grammar. (Foreshadowing: Even this much is not enough.)
- Relevant markedness constraints are adapted from Coetzee & Pater (submitted):
 

OCP-Place	* = two adjacent consonants with the same place of articulation.
OCP-Place-son	* = two adjacent coronals with the same PoA and value of [son].
OCP-Place-cont	* = two adjacent coronals with the same PoA and value of [cont].

- c. Assume [l] is [-cont].
- d. Stimuli are (empirically) ambiguous between /pl/ and /pr/, or between /tl/ and /tr/
  - ⇒ those are the two most-harmonic parses, and other candidate parses, such as /təl/, are eliminated by high-ranked constraints off-stage (e.g., perceptual faithfulness).

(15) /l/ is a worse candidate after /t\_/ than after /p\_/:

		Context-sensitive markedness constraints			Context-free constraints (examples):	
		OCP-PI-son	OCP-PI-cont	OCP-PI	Max-[+lat]	*[+lat]
[pl]	↑ /pl/					*
	↑ /pr/				*	
[tl]	↓ /tl/		*	*		*
	↓ /tr/				*	

There are fewer "l" responses in the [t\_] context than the [p\_] context because, for a fixed liquid stimulus, grammar is less favorable to the /l/ parse over its /r/ competitor in the /t\_/ context than in the /p\_/ context.

<b>pl</b>	<b>tl</b> OCP-PI-cont OCP-PI *[+lat]
<b>pr</b> Max-[+lat]	<b>tr</b> Max-[+lat]

(16) Source of bias: Context-free markedness and faithfulness constraints do not contribute to the contextual effect.

I assume that no context-sensitive faithfulness constraints are relevant (of which more later).

⇒ The between-context difference in bias is due only to *context-sensitive markedness constraints*.

(17) Size of bias: The tableau is categorical—the winner is *infinitely* better than the loser—whereas there is a lot of variation in the experimental results:

- a. Within-stimulus variability: A given stimulus may be judged "l" on 30% of trials, and "r" on 70% of trials.
- b. Between-stimulus variability: Changing the acoustical parameters changes the rate of "l" response.

Without modeling that variability, there is no way to predict the absolute magnitude of the bias effect. There are ways to do that (see Appendix), but all that is essential for this argument is that

- c. More-harmonic candidates are favored over less-harmonic ones, and
- d. The bias increases along with the difference in harmony (the number and rank of unshared violations).

- (18) Autonomy: The production grammar is never consulted as to whether the candidates are legal outputs or not.

### 3.2. Illegal parse vs. more-marked illegal parse

- (19) Autonomy  $\Rightarrow$  The perceptual grammar ought to cause bias when the stimulus is ambiguous between two parses which are *both* illegal (or *both* legal) in production.

Some evidence supporting this hypothesis: "More illegal" parses are more disfavored in perception relative to legal ones (Moreton 2002; Berent et al. 2006, 2007).

- (20) Example: English-speaking listeners misperceive CCVC nonwords as CəCVC when CC is illegal for sonority reasons. The worse the sonority profile, the likelier the misperception (Berent et al. 2006).

- (21) Modelling assumptions:

- a. Sonority scale following Parker 2002:239–240:

Segment class:	/ptk/	/bdg/	/fθsʃ/	/vðzʒ/	/mɪŋ/	/l/	/r/	/w/
Sonority index:	1	2	3	4	6	9	10	11

- b. Relevant markedness constraints, adapted from Baertsch (1998):

Dist= $n$  Give one \* for  $C_1C_2$  onset where sonority ( $C_2$ ) – sonority ( $C_1$ ) =  $n$ .

Fixed ranking: Dist= $-9$  » Dist= $-8$  » ... » Dist= $9$

- (22) Perceptual epenthesis is more likely into the "more illegal" cluster

		Context-sensitive markedness constraints				Context-free constraints (examples):	
		Dist= $-7$	Dist= $0$	Dist= $4$		Dep	*[+lat]
[bn]	↓	/bn/		*			
		/bən/				*	
[lb]	↓	/lb/	*				*
		/ləb/				*	*

<b>bn</b>	<b>lb</b>
Dist= $4$	Dist= $7$
*[+lat]	*[+lat]
<b>bən</b>	<b>ləb</b>
Dep	Dep

#### 4. "Analysis by synthesis"

- (23) Perceptual effects of native-language phonotactics have also been attributed to a bias in favor of *legal production outputs* (Polivanov 1931; Messer 1967; Massaro & Cohen 1983; Hallé et al. 1998; Brown & Matthews 2001; see also Brown & Hildum 1956; Dupoux et al. 1999, 2001; Moreton 2002).
- (24) ⇒ Perception involves determining whether each perceptual candidate is a possible output of production. I will call this the Analysis-by-Synthesis Hypothesis, after an early proposal by Stevens and Halle (1967, and references cited therein).
- (25) Modelling Analysis-by-Synthesis in an OT perception grammar:
- ~~Brute force: Introduce a constraint YYY that says: "Give one \* if the candidate is not a possible output of the perception grammar".~~ The inversion problem is not solved in general (Hale & Reiss 1997; Eisner 2002). (More in Appendix.)
  - Approximate "is a possible output" by "is a fixed point of the production grammar". I.e., assume that the production grammar has no chain shifts, so that [x] is a possible output if and only if /x/→[x].
  - Add a new constraint in the perception grammar:  
FIXPT: Give one \* to every candidate which differs from its own production output.  
(Cf. the production faithfulness constraints in Boersma (1997 and subsequently), which compare the production input with the *perceptual parse* of the candidate production output.)
  - FIXPT is ranked relatively high in the perception grammar (we'll come back to this later).
- (26) The examples discussed above are consistent with this hypothesis....

		Context-sensitive markedness constraints			Context-free constraints (examples):		
		FIXPT	Dist=-7	Dist=0	Dist=4	Dep	*[+lat]
[bn]	↓ /bn/	<i>bən</i>			*		
	/bən/					*	
[lb]	↓ /lb/	<i>ləb</i>	*				*
	/ləb/					*	*

<b>bn</b>	<b>lb</b>
FIXPT	FIXPT
Dist=4	Dist=7
*[+lat]	*[+lat]
<b>bən</b>	<b>ləb</b>
Dep	Dep

... but that's the most that can be said; there isn't any positive evidence *for* FIXPT, since it isn't needed to account for the pattern.

## 5. Experiment

### 5.1. Design

(27) Present stop-sonorant stimuli that are *four* ways ambiguous:

BD array	–or–	MN array
bl	dl	ml
bw	dw	nw

and measure bias as the dependency between the stop and sonorant decisions (Nearey 1990; Moreton 2002).

(28) OCP-PI-cont and OCP-PI contribute equally to the phonotactic effect in both arrays, as do the context-free markedness and faithfulness constraints.

⇒ Relative size of effects depends on the existence and ranking of OCP-PI-son and FIXPT.

		OCP constraints					Onset sonority constraints (ignore now)			
		FIXP T	OCP- PI-son	OCP- PI-cont	OCP-PI	Dist=3	Dist=5	Dist=7	Dist=9	
BD array	↑	/bl/						*		
		/bw/			*				*	
	↓	/dl/	dəl		*	*			*	
		/dw/							*	
MN array	↑	/ml/	məl			*				
		/mw/	məw	*		*		*		
	↓	/nl/	nəl	*	*	*	*			
		/nw/	nəw					*		

(29) Autonomous Perception Grammar hypothesis: No FIXPT, so MN effect should be at least as big as BD effect (since there is one more constraint, OCP-Pl-son, assisting).

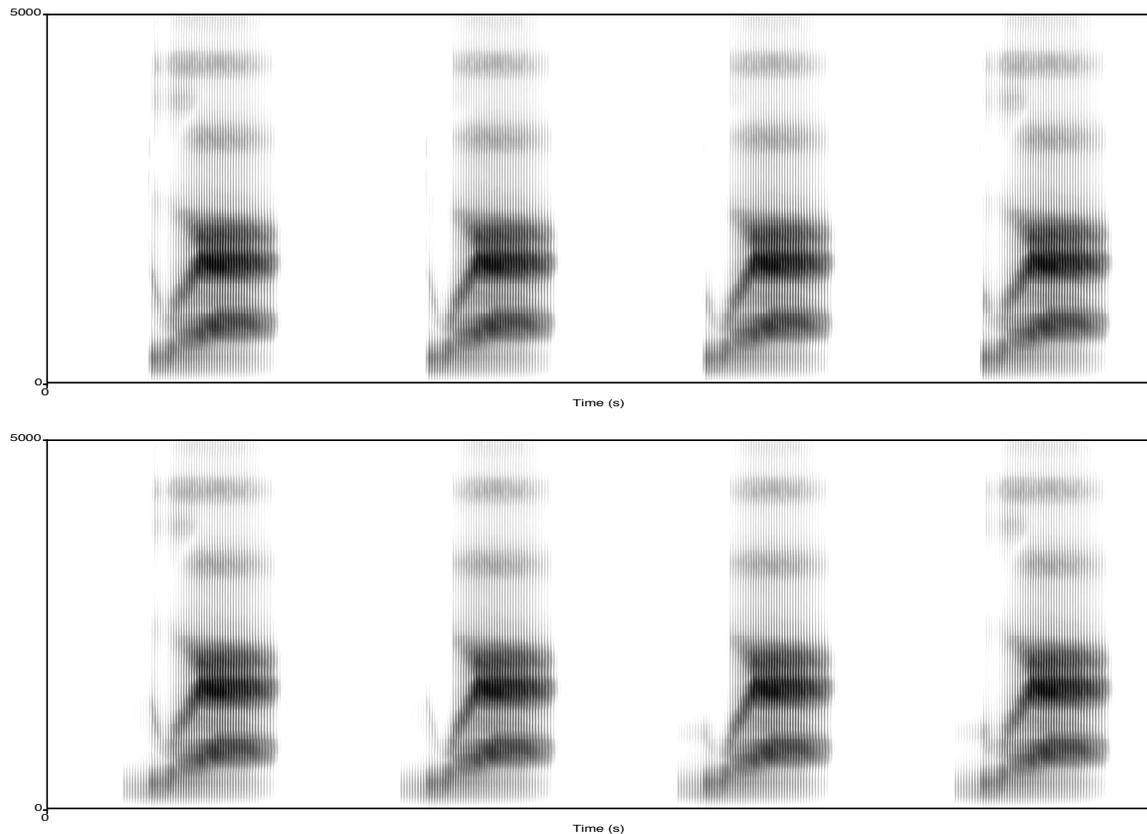
BD array		MN array	
<b>bl</b>	<b>dl</b> OCP-Pl OCP-Pl-cont	<b>ml</b>	<b>nl</b> OCP-Pl OCP-Pl-cont OCP-Pl-son
<b>bw</b> OCP-PL	<b>dw</b>	<b>mw</b> OCP-Pl OCP-Pl-son	<b>nw</b>

(30) Analysis-by-Synthesis hypothesis: If FIXPT is ranked far enough above OCP-Pl-son, it could happen that the MN effect is *smaller* than the BD effect.

BD array		MN array	
<b>bl</b>	<b>dl</b> FIXPT OCP-Pl OCP-Pl-cont	<b>ml</b> FIXPT	<b>nl</b> FIXPT OCP-Pl OCP-Pl-cont OCP-Pl-son
<b>bw</b> OCP-PL	<b>dw</b>	<b>mw</b> FIXPT OCP-Pl OCP-Pl-son	<b>nw</b> FIXPT

## 5.2. Stimuli

- (31) Two 6x6 arrays, synthesized using SENSYN (Sensimetrics Corporation) on the basis of the acoustical theory of the relevant segments (Fujimura 1962; Recasens 1983; Stevens 1999:307ff., 489ff., 523ff., 535ff) and the author's own productions analyzed in Praat (Boersma & Weenink 2002). The MN array was perfected first; then the BD array was made from it by replacing the nasal murmur with closure voicing. Corresponding MN and BD stimuli were identical after the release.



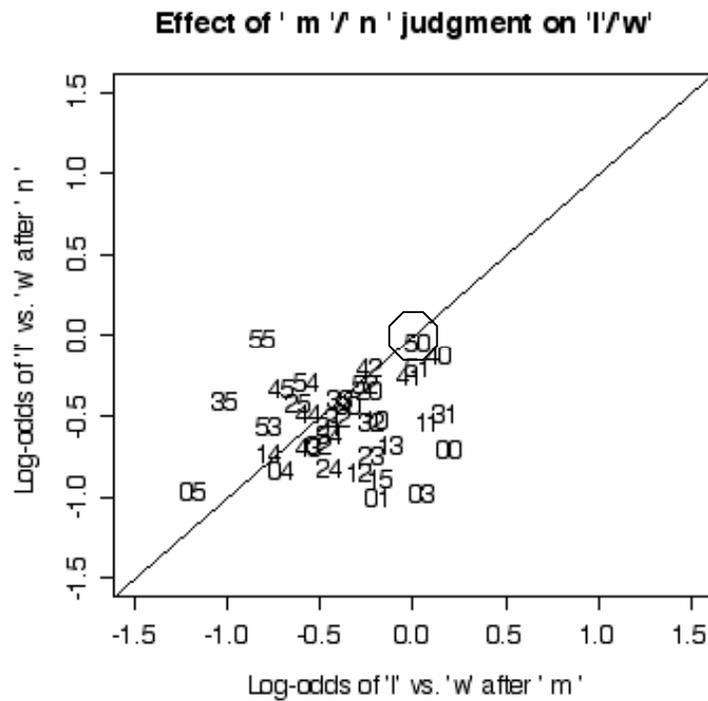
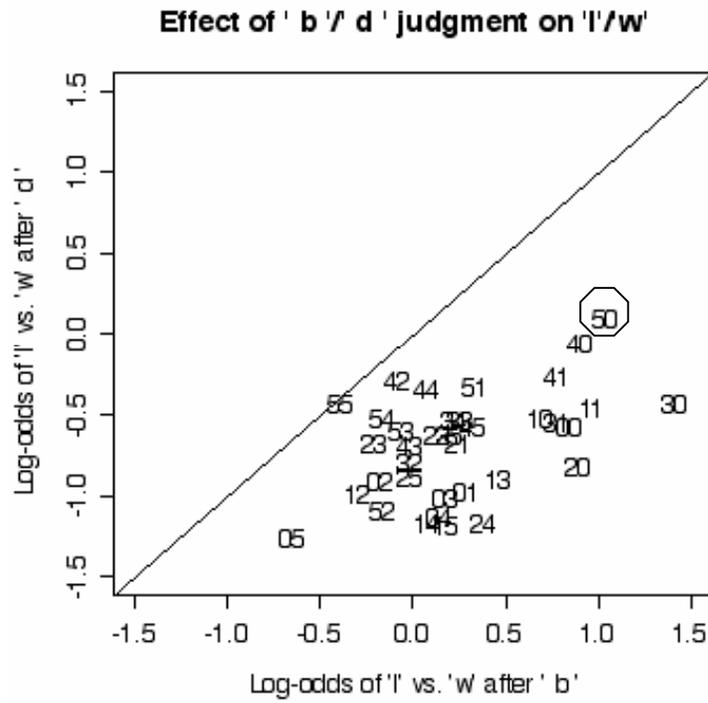
## 5.3. Participants and procedure

- (32) 16 native speakers of American English. Half did BD array first, half did MN array first.
- (33) Each array was presented separately through headphones in a quiet room 10 times for 4-choice identification, under control of a script written in MacPerl 5.6 (Neeracher & Nandor 2002). Buttons on screen were labelled with some permutation of “dl dw bl bw” or “nl nw ml mw” (same within each subject, counterbalanced across subjects). Eight listeners did the b/d array first, and eight did m/n first. Each filled out a post-experiment questionnaire on language background and subjective experience of the stimuli.

## 5.4. Results

- (34) For each stimulus, responses are pooled across listeners, and
- l/w judgment is quantified as log-odds of “l” vs. “w” responses to that stimulus,
  - computed separately for each stop response,
  - and plotted on the  $x$  and  $y$  axes.
- (35) No bias  $\Leftrightarrow$  l/w log-odds is the same regardless of stop response  $\Leftrightarrow x = y \Leftrightarrow$  point lies on the line  $y = x$ .

Hence, displacement from  $y = x$  indicates bias.



(36) Mixed-effects logistic-regression model with Subject as a random effect, Sonorant response as dependent variable (positive means more "w" responses). Independent variables:

- a. Stop            Position along [d]–[b] or [n–m] scale (0-5)
- b. Son             Position along [l]–[w] scale (0–5)
- c. MNarray       0 for BD array, 1 for MN array (i.e., BD array is reference category)
- d. CorResp       0 for "b" or "m" responses, 1 for "d" or "n" responses.

Interaction terms of interest:

- e. Stop:CorResp    Compensation for coarticulation (should be positive)
- f. CorResp:MNarray Difference in phonotactic effect between BD and MN array

Fixed effects	Value	StdErr	p-value	
Intercept	-0.7228	0.1255	0.0000	*** Mostly "l" after "b" response.
Stop	0.0129	0.0332	0.6982	Stop acoustics didn't affect "l" rate,
Son	0.1566	0.0312	0.0000	*** ... but sonorant acoustics did
CorResp	1.4577	0.2711	0.0000	*** More "w" after "d" response.
MNarray	0.7575	0.1900	0.0001	*** More "w" in MN array generally.
CorResp:MNarray	-0.8010	0.2593	0.0020	** Smaller phonotactic effect in MN
Son:MNarray	-0.0418	0.0432	0.3325	
Stop:CorResp	-0.1521	0.0526	0.0038	**

(37) ⇒ Bias in the MN array is reduced compared to the BD array, contrary to the predictions of the Autonomous Perception Grammar hypothesis.

## 6. Alternative explanations

(38) Stop decision affected sonorant decision more in the BD array than the MN array. Consistent with Analysis-by-Synthesis; inconsistent with Autonomous Perception Grammar, but are there any alternative explanations that rescue the Autonomous Perception Grammar hypothesis?

(39) Perceptual epenthesis. Maybe the Autonomous Perception Grammar hypothesis is right, but listeners were perceiving MN array with epenthesis ([məl], [məw], etc., as in Berent et al. 2006, 2007), making the syllable-onset OCP-Place constraints irrelevant.

- a. Not noticed by experimenter during stimulus construction (main problems were keeping [nw] from sounding like [lw], and [bw] from sounding like [gw])
- b. Not spontaneously reported during supervised, interactive practice phase at start of each array (corner stimuli repeated until all four response options had been used).
- c. First question on post-experiment questionnaire asked what they had heard besides what was on the buttons. Answers:
  - i. 7 reported hearing only what they were instructed to hear.
  - ii. 4 reported only changes that would make the MN and BD arrays more alike: nasals heard as oral stops or vice versa ("at one point").
  - iii. 3 reported percepts that would not have reduced the OCP-Place effect in the MN array relative to that in the BD array: [lw] ("maybe", "sometimes", "4–5"), [p] for [b] ("feel like").
  - iv. Only 3 reported misperceptions that might have reduced the MN effect ([w] or [m] for [ml]/[mw], [mbw] for [mw], [l] for [ml]/[nl]).

⇒ possible, but unlikely.

- (40) Markedness swamping. Maybe the Autonomous Perception Grammar hypothesis is still right, but some onset-sonority constraints are ranked so high that they swamped the effect of the OCP constraints:

			Dist=3	Dist=5	OCP-PI-son	OCP-PI-cont	OCP-PI	Dist=7	Dist=9
BD array	↑	/bl/						*	
		/bw/					*		*
	↓	/dl/				*	*	*	
		/dw/							*
MN array	↓	/ml/	*						
		/mw/		*	*		*		
	↓	/nl/	*		*	*	*		
		/nw/		*					

- The argument is that Dist=3 strongly favors the [w] parse in the MN array regardless of the stop decision, and is so high-ranked that the three OCP-PI constraints are ineffectual in comparison.
- Indeed, the rate of "w" response is higher in the MN array than the BD array (the MN array coefficient is 0.76 logits), suggesting that Dist=3 is making itself felt.
- However, this effect of Dist=3 is *smaller* than that of the OCP constraints in the BD array (the CorResp coefficient is 1.46 logits).
- Last-ditch effort: Maybe it's a range or anchoring effect, such that perceptual bias is determined by comparison with the most-marked response option; /dl/ being much worse than the others in the BD array, but /ml nl/ being about equally bad compared to /mw nw/. But then we should see large perceptual-bias effects between *legal* stimuli as well, which doesn't seem to happen (Brown & Hildum 1956; Messer 1967; Berent et al. 2006.)

(41) Context-sensitive faithfulness. Perhaps these cannot be safely ignored, contra (16) above. Steriade (2001, 2002) has proposed that there are context-sensitive production-faithfulness constraints ranked so as to inhibit candidate outputs in proportion to their auditory distinctness from the input.

a. Assuming such constraints exist in perception as well, under the right circumstances, they could in principle overcome the OCP effects:

			Max-[+lat]/ n_	OCP constraints			Max-[+lat]/ d_
				OCP- PI-son	OCP- PI-cont	OCP-PI	
BD array	↑	/bl/					
		/bw/			*		
	↓	/dl/		*	*		
		/dw/				*	
MN array	↑	/ml/					
		/mw/		*	*		
	↓	/nl/		*	*	*	
		/nw/	*				

b. The general prediction is that greater auditory similarity → lower-ranked faithfulness → larger phonotactic effect, where auditory similarity is measured by (inter alia) confusability.

c. The plots show that the MN array is the more confusable one.

d. Also, the effect of changing the sonorant acoustics (a measure of perceptual faithfulness) does not differ between the two arrays (Son : MN array coefficient is near zero and not significant).

(42) ⇒ The experiment yielded pretty good evidence that production legality affects perception (Analysis-by-Synthesis).

## 7. Discussion

(43) Summary:

- Production legality is needed to account for the experimental results; the production constraints alone cannot. ⇒ Perception must be able to consult production grammar.
- On the other hand, production legality cannot account for *all* results of this experiment—there was still a phonotactic effect in the MN array, even though all parses were illegal in production. ⇒ Perception is not reducible to production grammar.
- Interaction between production legality and perception constraints can be modelled using the FIXPT constraint, which penalizes perception candidates that are changed by the production grammar. (Assumes production grammar has no chain shifts.)

(44) Analysis-by-Synthesis account of experiment requires FIXPT to be ranked very high (above unviolated markedness constraints). Nevertheless, this ranking does not have to be stipulated:

- When production is immature, obeying FIXPT *causes* perceptual errors (two different adult words are perceived as the same). ⇒ Error-driven learning (reviewed in, e.g., Tessier 2007) keeps FIXPT low.

- b. When production is adult-like, obeying FIXPT *prevents* perceptual errors. ⇒ Error-driven learning moves it up.
  - c. ⇒ Prediction: Production illegality should matter less to children than to adults in perception tasks. (I.e., children should mistrust their own production as a guide to perception.)
- (45) Allows modelling of task effects (modulation of phonotactic effect depending on experimental situation) as reranking of FIXPT (e.g., Hallé et al. 1998; Berent et al. 2006) rather than as shifting of attention between representational levels. This makes possible unified analyses of hyperarticulation and "hyperperception".
- (46) "I asked you first" Implications for theories in which production consults perception:
- a. The H&H Theory (Lindblom 1990, 1996): Speaker controls articulatory effort in real time so as to provide "*sufficient contrast*, that is, discriminative power that is sufficient for lexical access" (1990:405), while minimizing articulatory effort. To do this, speaker uses internal model of the non-signal information available to the listener, including linguistic knowledge.
  - b. Faithfulness in Functional Phonology (Boersma 1997a, 1998:143–148, 2001, 2005): The articulatory production candidates are used to generate input to the perception grammar, and the resulting percept is compared with the lexical representation.
  - c. The P-map (Steriade 2001, 2002): The perceptibility model ("P-map") which determines the production-grammar faithfulness rankings is partially language-dependent, meaning that the perception grammar (not discussed by Steriade) would have to be consulted in order to set up the production grammar.

On-line use of FIXPT in perception is incompatible with on-line use of the perception grammar in production, since one or the other must return an answer first. (Unless one or both is caching outputs.)

---

## Appendix

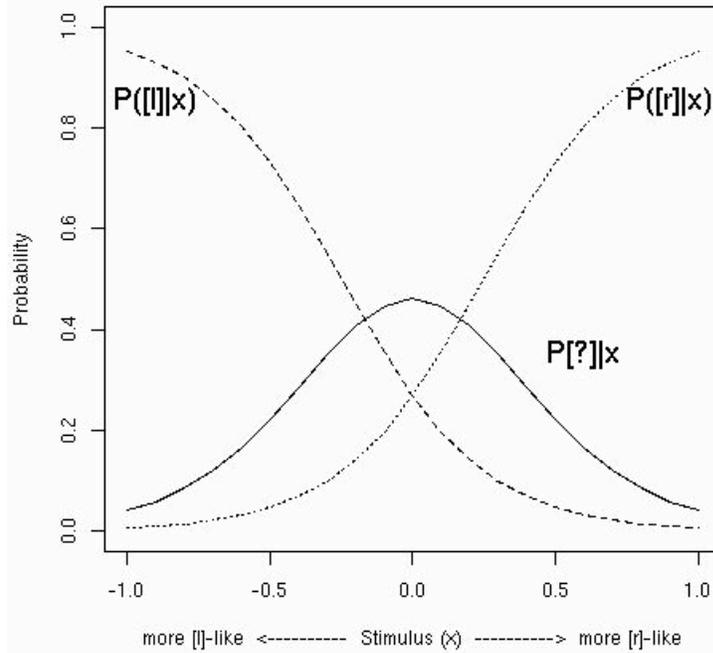
(A1) Modelling within-stimulus variation in perception: Multiple repetitions of the same stimulus can evoke different percepts in the same perceiver. This is analogous to free variation in production, and might be modelled using the same techniques, such as partial ranking (Antttila 1995; Anttila & Cho 1998), floating constraints (Reynolds 1994:115–119, Nagy & Reynolds 1997), and stochastic ranking (Boersma 1997b, 1998, Boersma & Hayes 2001; Zubritskaya 1997). Alternatively, a squashing function might be used to convert harmony to a scalar quantity (e.g., by weighting the constraints), and the probability of a particular parse could then be determined by the Luce choice rule (Luce 1959).

However, some of what appears to be within-stimulus variation may actually be *between*-stimulus variation, as the same acoustic stimulus may give rise to different early auditory representations on different trials.

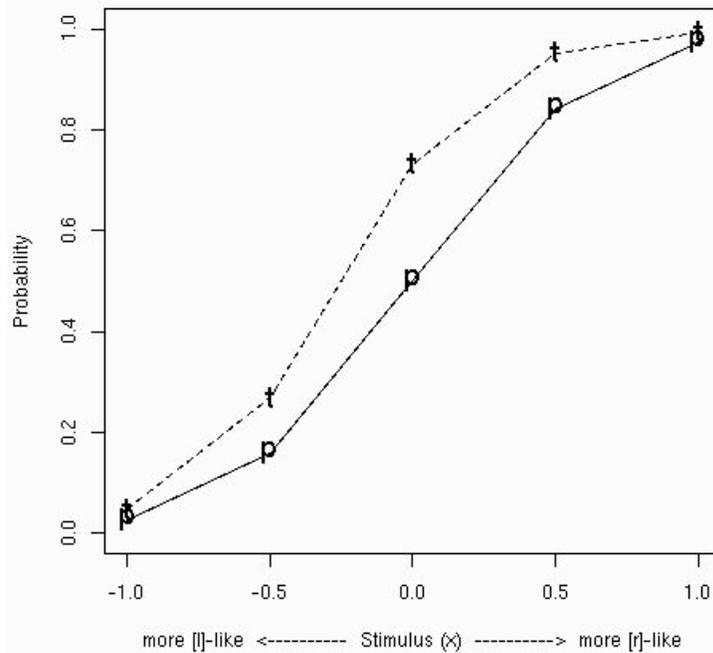
(A2) Modelling between-stimulus variation in perception: Whence the S-shaped curves in the Massaro-Cohen results? Here is a simple model which does not sacrifice discreteness in the phonological grammar (inspired by Macmillan & Creelman 1990, Chapter 4):

- a. At the phonetic level, a given acoustic stimulus evokes the phonetic representation [l] with probability  $p$ , [r] with probability  $q$ , and [?]—a featurally intermediate representation, such as a liquid underspecified for a binary [lateral] feature—with probability  $1-p-q$ . Both  $p$  and  $q$  depend on the acoustics of the stimulus.
- b. The grammar then operates on the phonetic representation to yield a phonological parse. Suppose the constraints are ranked such that phonetic [l] and [r] are invariably perceived correctly as /l/ and /r/; phonetic [?] is perceived as /l/ with probability  $s$  and as /r/ with probability  $1-s$ . These probabilities are independent of the acoustics of the stimulus.
- c. The probability that a given stimulus  $x$  will be perceived as /l/ is therefore
$$P(/l/ | x) = p + (1-p-q)s$$
- d. Example: Suppose  $x$  ranges from  $-1$  (most [r]-like) to  $+1$  (most [l]-like), and say that
  - $p = P([l] | x) = \exp(4x - 1) / (1 + \exp(4x - 1))$
  - $q = P([r] | x) = \exp(-(4x - 1)) / (1 + \exp(-(4x - 1)))$
  - $s = P(/l/ | [?]) = 1/2$  in the context [p\_],  $0$  in the context [t\_] (i.e., the ranking in the perceptual grammar is such that the intermediate phonetic representation has a 50% chance of being mapped to [l] and [r] in the [p\_] context, but is always mapped to [r] in the [t] context.

e. Changing the stimulus changes the probabilities of the three phonetic representations...



f. ... which in turn changes the probability of getting an /r/ response (cf. 13c above)



(A3) What if the grammar has a chain shift? Then being a possible production output is no longer equivalent to being a fixed point of the production grammar, and FIXPT will predict perceptual bias against the middle member of an  $a \rightarrow b \rightarrow c$  chain shift. This bias may or may not exist. If it does not, and the correct generalization is that the bias applies only to possible outputs, we are then faced with the problem of determining what a possible output is in a grammar with chain shifts. This is closely related to the problem of inverting the grammar. Hale and Reiss (1997) have proposed two algorithms for doing this in the general case; the first is brute force (try every possible input); the second starts with a non-empty set of potential inverted inputs and continually enlarges it, and hence cannot ever answer "no".

A more focused search is possible if one adopts the hypothesis that chain shifts arise through local conjunction of faithfulness constraints (Kirchner 1997).

- a. Given a form [b], such that we want to know whether it is a possible output, we first try it as an input, and see if  $b \rightarrow b$ . If so, we are done.
- b. Otherwise,  $b \rightarrow c$  for some  $c \neq b$ . Then [b] is only a possible output if there is some  $a \neq b$  such that  $a \rightarrow b \rightarrow c$ .
- c. If  $a \rightarrow b \rightarrow c$ , then there must be a faithfulness constraint F such that  $F/a/[c] > F/b/[c] = 0$  (Prince 1998, Cable 2000, Moreton & Smolensky 2002). There may be a lot of faithfulness constraints satisfying that condition.
- d. If Kirchner (1997) is right, then that F must be the conjunction of two other constraints, F1 and F2, such that  $a \rightarrow b$  violates F1 and  $b \rightarrow c$  violates F2. So we narrow our focus to conjoined faithfulness constraints meeting this description.
- e. Since we know what b and c are, we know which of the conjuncts is F2 (it's the one violated by  $b \rightarrow c$ ). Hence, we know which conjunct is F1 (it's the other one).
- f. So now we know that if there is such an /a/, then  $a \rightarrow b$  has to violate F1. Our search space is now narrowed to the set to /a/ fitting that description.

Even this reduced search space may still be infinite, which is no improvement over the brute-force algorithm. I.e., the problem is still not solved.

---

## References

- Anttila, Arto. 1995. Deriving variation from grammar: a study of Finnish genitives. MS, Stanford University.
- Anttila, Arto, and Young-Mee Yu Cho. 1998. Variation and change in Optimality Theory. *Lingua* 104:31–56.
- Baertsch, Karen. 1998. Onset sonority distance constraints through local conjunction. Papers from the 34th Meeting of the Chicago Linguistics Society, Part 2: Papers from the Panels, 1–15.
- Berent, Iris, Donca Steriade, Tracy Lennertz, and Vered Vaknin. 2006. What we know about what we have never heard: evidence from perceptual illusions. To appear in *Cognition*. *Cognition* 2006, doi:10.1016/j.cognition.2006.05.015.
- Berent, Iris, Donca Steriade, and Tracy Lennertz. 2007. Speakers' sensitivity of the markedness of unattested onsets. Presentation at the 26th West Coast Conference on Formal Linguistics, April 27–29, Berkeley.
- Boersma, Paul. 1997a. Functional Optimality Theory. *Proceedings of the Institute of Phonetic Sciences of University of Amsterdam* 21:37–42.
- Boersma, Paul. 1997b. How we learn variation, optionality, and probability. MS, Rutgers University: Rutgers Optimality Archive.
- Boersma, Paul. 1998. Functional Phonology: formalizing the interactions between articulatory and perceptual drives. Ph.D. dissertation, University of Amsterdam. The Hague: Holland Academic Graphics.

- Boersma, Paul. 1999. On the need for a separate perception grammar. MS, Rutgers University: Rutgers Optimality Archive.
- Boersma, Paul. 2000. The OCP in the perception grammar. MS, Rutgers University: Rutgers Optimality Archive.
- Boersma, Paul. 2001. Phonology-semantics interaction in OT, and its acquisition. In: Robert Kirchner, Wolf Wikeley, and Joe Pater (eds.), *Papers in Experimental and Theoretical Linguistics 6: Workshop on the Lexicon in Phonetics and Phonology*, 24–35. Edmonton: University of Alberta.
- Boersma, Paul. 2005. Some listener-oriented accounts of *h*-aspiré in French. MS, Rutgers University: Rutgers Optimality Archive.
- Boersma, Paul, and Silke Hamann. 2007. The evolution of auditory contrast. MS, Rutgers University: Rutgers Optimality Archive.
- Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32(1):45–86.
- Braine, Martin D. S. 1976. Review of Neilson V. Smith, *The acquisition of phonology: A case study*. *Language* 52(2):489–498.
- Brown, Cynthia, and John Matthews. 2001. When intake exceeds input: language specific perceptual illusions induced by L1 prosodic constraints.
- Brown, Roger W., and Donald C. Hildum. 1956. Expectancy and the perception of syllables. *Language* 32:411–419.
- Cable, Seth. 2000. Constraint rankings that map only onto themselves. MS, Rutgers University.
- Coetzee, Andries, and Joe Pater. (submitted). Lexically ranked OCP-Place constraints in Muna. MS, University of Michigan.
- Donegan, Patricia. 2001. Constraints and processes in phonological perception. In: Katarina Dziubalska-Kolaczyk (ed.), *Trends in Linguistics: Constraints and Preferences*, 43–68.
- Dupoux, E., K. Kakehi, Y. Hirose, C. Pallier, and J. Mehler. Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25(6):1568–1578.
- Eisner, Jason. 2002. Comprehension and compilation in Optimality Theory. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 56–63.
- Fujimura, Osamu. 1962. Analysis of nasal consonants. *Journal of the Acoustical Society of America* 34(12):1865–1875.
- Hale, Mark, and Charles Reiss. 1997. How to parse (and how not to) in OT phonology. *Papers from the 27th Meeting of the North-East Linguistics Society*, 159–170. Amherst, Mass.: Graduate Linguistic Students' Association.
- Hallé, Pierre A., Juan Segui, Uli Frauenfelder, and Christine Meunier. 1998. Processing of illegal consonant clusters: a case of perceptual assimilation? *Journal of Experimental Psychology: Human Perception and Performance* 24 (2):592–608.
- Ingram, David. 1974. Phonological rules in young children. *Journal of Child Language* 1:49–64.
- Kabak, Baris, and William Idsardi. 2003. Syllabically conditioned perceptual epenthesis. In: Pawel M. Nowak, Corey Yoquelet, and David Mortensen (eds.), *Proceedings of the 29th annual meeting of the Berkeley Linguistics Society*, February 14–17.
- Kenstowicz, Michael. 2001. The role of perception in loanword phonology. *Linguistique Africaine* 20:??–??.
- Kirchner, Robert. 1996. Cues or contexts in feature licensing constraints. MS, Rutgers University: Rutgers Optimality Archive.
- Kirchner, Robert. 1997. Synchronic chain shifts in Optimality Theory. *Linguistic Inquiry* 27(2):341–350.
- Lassette, Paul, and Patricia Donegan. 1998. Perception in Optimality Theory: the frugality of the base. *Papers from the 24th Meeting of the Berkeley Linguistics Society*, 346–355.
- Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In: W. J. Hardcastle and A. Marchal (eds.), *Speech Production and Speech Modelling*, 403–439. Dordrecht: Kluwer.
- Lindblom, Björn. 1996. Role of articulation in speech perception: clues from production. *Journal of the Acoustical Society of America* 99 (3):1683–1692.

- Luce, R. Duncan. 1959 [2005]. *Individual choice behavior: a theoretical analysis*. Mineola: Dover.
- Macken, Marlys A. 1980. The child's lexical representation: the 'puzzle-puddle-pickle' evidence. *Journal of Linguistics* 16:1-17.
- Macmillan, Neil A., and C. Douglas Creelman. 1990. *Detection Theory: A User's Guide*. Cambridge, UK: Cambridge University Press.
- Massaro, Dominic W., and Michael M. Cohen. 1983. Phonological context in speech perception. *Perception and Psychophysics* 34(4):338–348.
- Messer, Stanley. 1967. Implicit phonology in children. *Journal of Verbal Learning and Verbal Behavior* 6:609–613.
- Miller, George A., and Noam Chomsky. 1963. Finitary models of language users. In: R. Duncan Luce, Robert R. Bush, and Eugene Galanter (eds.), *Handbook of Mathematical Psychology, Volume II*, 419–492.
- Moreton, Elliott (2002). Structural constraints in the perception of English stop-sonorant clusters. *Cognition* 84:55–71.
- Moreton, Elliott, and Shigeaki Amano (1999). Phonotactics in the perception of Japanese vowel length: Evidence for long-distance dependencies. *Proceedings of the 6th European Conference on Speech Communication and Technology, Budapest*
- Moreton, Elliott, and Paul Smolensky. 2002. Typological consequences of local constraint conjunction. *Proceedings of the 21st meeting of the West Coast Conference on Formal Linguistics*.
- Nagy, Naomi, and William Reynolds. 1997. Optimality Theory and variable word-final deletion in Faetar. *Language Variation and Change* 9(1):37–56.
- Nearey, Terrence M. 1990. The segment as a unit of speech perception. *Journal of Phonetics* 18:347–373.
- Ohala, Diane. 1996. Cluster reduction and constraints in acquisition. PhD dissertation, U. of Arizona.
- Pater, Joe. 1998. From phonological typology to the development of receptive and productive phonological competence: applications of minimal violation. MS, Rutgers University: Rutgers Optimality Archive.
- Parker, Stephen G. 2002. Quantifying the Sonority Hierarchy. Ph.D. dissertation, University of Massachusetts, Amherst.
- Pisoni, David B., and Paul A. Luce. 1987. Acoustic-phonetic representations in word recognition. *Cognition* 25:21–52.
- Pitt, Mark A. 1998. Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception and Psychophysics* 60:941–951.
- Polivanov, Evgenij. La perception des sons d'une langue étrangère. *Travaux du Cercle linguistique de Prague*.4:79–96.
- Prince, Alan. 1998. Current directions in Optimality Theory. Handout from talk presented at the Phonology Forum, Kobe University, September 4. Rutgers University: Rutgers Optimality Archive.
- Recasens, Daniel. 1983. Place cues for nasal consonants with special reference to Catalan. *Journal of the Acoustical Society of America* 73(4):1346–1353.
- Reynolds, William T. 1994. Variation and phonological theory. Ph.D. dissertation, University of Pennsylvania.
- Silverman, Daniel. 1992. Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 9(2):289–328.
- Smolensky, Paul. 1996. On the comprehension/production dilemma in child language. *Linguistic Inquiry* 27:720–731.
- Steriade, Donca. 2001. Directional asymmetries in place assimilation: a perceptual account. In: Elizabeth Hume and Keith Johnson (eds.), *The Role of Speech Perception in Phonology*, 219–250. San Diego: Academic Press.
- Steriade, Donca. 2002 (to appear). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. MS, 11/12/02. (To appear in Kiparsky festschrift.)

- Stevens, Kenneth N., and Morris Halle. 1967. Remarks on analysis by synthesis and distinctive features. In: Weiant Wathen-Dunn (ed.), *Models for the Perception of Speech and Visual Form*, 88–102. Cambridge, Mass.: MIT Press.
- Stevens, Kenneth N. 1999. *Acoustic Phonetics*. Cambridge, Mass.: MIT Press.
- Studdert-Kennedy, Michael. 1974. The perception of speech. In: Thomas A. Sebeok (ed.), *Current Trends in Linguistics* 12:2349–2385.
- Tessier, Anne-Michelle. 2007. Biases and stages in phonological acquisition. Ph.D. dissertation, University of Massachusetts, Amherst.
- Zubritskaya, Katya. 1997. Mechanism of sound change in Optimality Theory. *Language Variation and Change* 9:121–148.