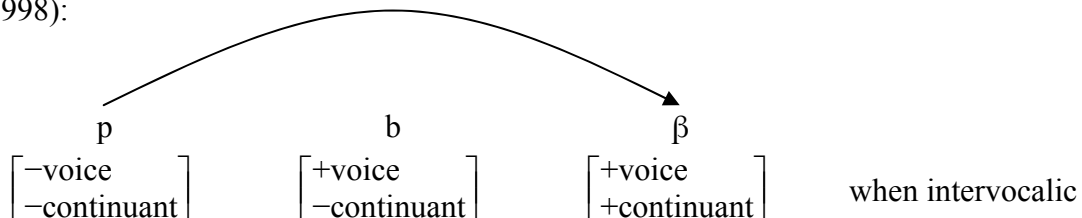# Saltation in Phonology[1]

## I. OVERVIEW

## 1. Defining saltation

- Term proposed by Lass (1997)[2]
- Intuitively: one sound "leaping over" another. E.g. Campidanian Sardinian (Bolognesi 1998):



p
$\begin{bmatrix} -\text{voice} \\ -\text{continuant} \end{bmatrix}$

b
$\begin{bmatrix} +\text{voice} \\ -\text{continuant} \end{bmatrix}$

β
$\begin{bmatrix} +\text{voice} \\ +\text{continuant} \end{bmatrix}$      when intervocalic

## 2. Campidanian examples (Bolognesi 1998: 30-31, 36-39)

a. Intervocalic spirantization of voiceless stops

| | | |
|---|---|---|
| beːu [p]iʃːi | → beːu [β]iʃːi | 'nice fish' |
| sːu [t]rintaduzu | → sːu [ð]rintaduzu | 'the thirty-two' |
| dɛ [k]uatːru | → dɛ [ɣ]uatːru | 'of four...' |

b. Retention of intervocalic /b, d, g/

| | | |
|---|---|---|
| sːu [b]inu | → sːu [b]ĩu | 'the wine' |
| donːia [d]ominiɣu | → donːja [d]ominiɣu | 'every Sunday' |
| dɛ [g]ɔma | → dɛ [g]ɔma | 'of rubber' |

## 3. This talk: an empirical/analytic/theoretical tour of saltation

- Examples
  - ➢ Themes: apparent rarity, unusual diachronic origins

- Analysis
  - ➢ why classical Optimality Theory fails
  - ➢ why the Lubowicz analysis is not a good idea
  - ➢ A new proposal, based on Zuraw's *MAP and Steriade's P-map

---

[1] This represents a paper in progress written with James White of UCLA.

[2] We are avoiding Lubowicz's term for saltation, "derived environment effect": it's tied to a particular analysis, and we would prefer to use an analysis-independent label for the phonemenon.

- Psycholinguistics
  - ➢ Saltation appears to be hard to learn — work of White

- Learnability
  - ➢ Computational work explaining difficulty of learning saltation — work of White

- What it all means for phonology and phonological inquiry

## II. SALTATION EXAMPLES AND THEIR DIACHRONIC ORIGIN

**4. Theme**

- We cover a few cases of saltation (we only know of about ten)
- Where we know the history, result is always the same:  saltation is always a **restructured** process — never a fresh sound change.
- This in fact follows from earlier claims by Minkova (1993) and Lass (1997) re. gradualness of sound change — if you can't saltate across a *possible* segment, then you certainly can't saltate to create an alternation.

**5.  Type I:  A → C; B is later interposed in loanwords**

- **Standard German** (Ito and Mester 2003)
  - ➢ /g/ saltates over /k/ to become [ç] / ɪ ___ ]ᵂₒᵣd

    /ˈkøːnɪg/ → [ˈkøːnɪç] 'king'    cf. [ˈkøːnɪgə] 'kings'

  - ➢ The invariant /k/ words are cosmopolitan forms like *Plastik* and *Batik.*

- **Other**
  - ➢ Slovak [e, o] saltate over [eː, oː] — borrowed phonemes — to arrive at [ie, uo] (Rubach 1991, Lubowicz 2012)
  - ➢ Polish [g] saltates over borrowed [dʒ] to arrive at [ʒ] (Lubowicz 2002)

**6.  Type II:  A → C, B is later interposed through grammar change**

- We think Campidanian was not always saltatory.
- It was originally a **lenition chain**:  [p] and [b] move toward null:

$$\left\{ \begin{array}{ccccc} p & \rightarrow & b & \rightarrow & \beta \\ b & \rightarrow & \beta & \rightarrow & \varnothing \end{array} \right\} \text{ / V \_\_ V}$$

- Schematic evolution:

| | *'nice fish'* | *'the wine'* | |
|---|---|---|---|
| a. | bɛlːu [p]iʃːi | sːu [b]ĩu | primal stage |
| b. | bɛlːu [b]iʃːi | sːu [β]ĩu | intermediate stage |
| c. | bɛlːu [β]iʃːi | sːu [∅]ĩu | crisis stage — massive stem-initial neutralization! |

- When /b/ (also [d, g]) reached null, this triggered **massive restructuring** — restoration of the isolation form in intervocalic contexts.

    d.  bɛlːu [β]iʃːi      sːu [b]ĩu                    post-crisis stage — restructuring

## 7.  Why think that this account is true?

- In neighboring dialects, stage (b) above is attested; [b] → [β] / V ___ V[3]
- In modern Campidanian, some [b] ~ ∅ alternations **remain as relics**:

    ➤  Alternation is optional, found only in some words ([sːa bia ~ sːa ia] 'the road')
    ➤  ∅ is employed in *careful* speech (Bolognesi p. 36-7), not the expected rapid speech (lenition).
    ➤  The eligible words appear to be frequent, core vocabulary.[4]
    ➤  All this suggests:  in careful speech, speakers are retrieving lexically listed allomorphs —relics of language history.

## 8.  Upshot

- Saltation was not a sound change in Campidanian; rather [b] was interposed between [p] and [β] due to the loss of productive [b] ~ ∅ alternation.

## 9.  Type III:  retreat by hypercorrection in North German

- Sources:
    ➤  for saltation:  Ito and Mester (2003)
    ➤  historical background:  Zhirmunskii (1962), Robinson (2000), Pilch (1996)

- Northern German:  the history of voiced velars via sound change:

| *frag-en ~ frag* | *dick-e ~ dick* | *mach-en ~ mach*[5] | Alternating pairs with \*g, \*k, \*x |
|---|---|---|---|
| g ~ g | k ~ k | x ~ x | pre-North German |
| ɣ ~ ɣ | k ~ k | x ~ x | Spirantization of /g/ after vowels |
| ɣ ~ **x** | k ~ k | x ~ x | Final Devoicing[6] |

---

[3] See Bolognesi (1998: 36), citing Virdis.

[4] Alternating forms have glosses like 'of', 'want', 'road', 'house'.  Non-alternating forms have glosses like 'doctor', 'drill', 'rubber', 'chicory'.

[5] These forms are schematic and not guaranteed to match actual dialect data.  Glosses:  'ask-imperative/infinitive'; 'fat-plain form/adjectival inflection'; 'do-imper./infin'.
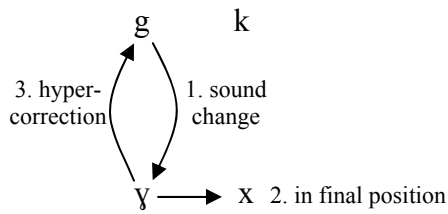
[6] There are complications to this history (Kiparsky (1982:39-40) which I ignore here.

- This is Vernacular Northern German.  In Refined Northern German:  [ɣ] restores to [g], imitating the standard

    **g ~ x**          k ~ k          x ~ x          hypercorrection

    Now, in /frag/ → [frax], /g/ saltates over /k/ in going to [x].

- Some speakers have [ɣ ~ x] in informal speech, saltatory [g ~ x] in careful speech.
- Summary diagram:



## 10.  Summing up the empirical picture for saltation[7]

- Cases are not abundant, we think.
- … and we would expect this if they only arise under exceptional circumstances, as above
- We will suggest later that saltation is hard to learn; perhaps also relevant…

<div align="center">III. ANALYSIS OF SALTATION IN OPTIMALITY THEORY</div>

## 11.  Saltation and Optimality Theory

- Classical OT (Prince and Smolensky 1993) is famous for needing repair to cover phonological **opacity**,[8] and there is a huge literature on the topic.

    ➢ Many different ways to augment classical OT to accommodate opacity.
    ➢ More on this at end of talk.

- But saltation is also problematic for classical OT:  Lubowicz (2002), Ito/Mester (2003).

## 12.  Why classical OT fails for saltation, I:  basic constraints and ranking for Campidanian

- Markedness constraints

    ➢ *V[−voice]V          *favors intervocalic voicing*
    ➢ *V[−cont]V          *favors intervocalic spirantization*

---

[7] The remaining empirical examples known to us occur in Manga Kanuri (Schuh, ms.), Suma (Bradshaw 1999), and Japanese (Ito and Mester 1997). If you are aware of any more please let us know!

[8] See McCarthy (2007) and Baković (2007) for comprehensive overviews. The two fundamental cases, expressed in rules, are:  (i) A → B / C ___D, then new CAD created by a second rule; (ii) A → B / C ___D, then a later rule removes or crucially alters CBD.

- Faithfulness constraints

  ➢ IDENT(voice)          *favors retention of basic voicing value*
  ➢ IDENT(cont)           *favors retention of basic stop status*

- We get /p/ → [β] only if each Markedness constraint dominates its opposing Faithfulness constraint:

| /apa/ | *V[−voice]V | * V[−cont]V | IDENT(−voice) | IDENT(cont) |
|---|---|---|---|---|
| ☞ aβa |  |  | * | * |
| aba |  | *! | * |  |
| aɸa | *! |  |  | * |
| apa | *! | * |  |  |

## 13. Why classical OT fails here

- Once we've set up the grammar this way, /b/ is doomed to spirantize too:

| /aba/ | *V[−voice]V | * V[−cont]V | IDENT(−voice) | IDENT(cont) |
|---|---|---|---|---|
| ☞ *aβa |  |  |  | * |
| aba |  | *! |  |  |
| aɸa | *! |  | * | * |
| apa | *! | * | * |  |

- The same sort of argument applies for saltation in general.

## 14. A remedy for this problem:  Lubowicz (2002), Ito and Mester (2003)

- **Local constraint conjunction** (Smolensky 1995)
- This idea has a decent pedigree, e.g. in conjoining markedness constraints (Ito/Mester):

  ➢ It's bad for obstruents to be voiced.
  ➢ It's bad for consonants to be in coda position.
  ➢ It's especially bad to be a voiced obstruent coda
  ➢ Local conjunction:  *VOICED OBSTRUENT & *CODA (within same segment)

- Lubowicz's idea for saltation:  *conjoin Markedness and Faithfulness*.

## 15. The necessary conjoined constraint for Campidanian

- IDENT(voice) & *V[−cont]V
  ➢ "Don't both be unfaithful to voicing and be an intervocalic stop."

*a. Conjoined constraint forces /apa/ to become* [aβa]

| /apa/ | **ID(vce) & *V[−cont]V** | *V[−vce]V | ID(voice) | ID(cont) | *V[−cont]V |
|---|---|---|---|---|---|
| ☞ aβa | | | * | * | |
| *apa | | *! | | | * |
| *aba | *! | | * | | * |

> ➢ Without the conjoined constraint, *[aba] would win.

*b. /aba/ is stable*

| /aba/ | ID(vce) & *V[−cont]V | *V[−vce]V | ID(voice) | ID(cont) | *V[−cont]V |
|---|---|---|---|---|---|
| ☞ aba | | | | | * |
| *aβa | | | | *! | |

- This approaches generalizes to all cases of saltation.

**16.  This solves one problem, but at the same time creates a worse one[9]**

- Conjoined Markedness & Faithfulness produces disastrous predictions when applied to **phonotactics** (inventory of legal sequences).

**17.  Let's derive this language:  "voiced obstruents are legal only in triple consonant clusters"**

- The **devil's constraint**:  IDENT(voice) & *CCC
- Backdrop:  a language like Hawaiian with (almost) no voiced obstruents
  > ➢ Thus:  *[−son, +voice] >> IDENT(voice)

- But our version of Hawaiian will include the devil's constraint, ranked highest.

*a. Voiced obstruents disallowed in simple cases*

| /ba/ | IDENT(voice) & *CCC | MAX(C) | *[−son, +voice] | *CCC | IDENT(voice) |
|---|---|---|---|---|---|
| ☞ pa | | | | | * |
| **\*ba** | | | *! | | |

---

[9] This section of the talk channels a personal communication from Colin Wilson.

*b. Voiced obstruents allowed in triple clusters*

| /atbka/ | IDENT(voice) & *CCC | MAX(C) | *[–son, +voice] | *CCC | IDENT(voice) |
|---|---|---|---|---|---|
| ☞ **atbka** | | | * | * | |
| *atka | | *! | | | |
| *atpka | *! | | | * | * |

- This bad prediction generalizes:  something marked, and generally illegal, *suddenly becomes ok when it's in an overlapping marked context*.

## IV. LOOKING FOR A BETTER ANALYSIS

**18. Desiderata**

- Expand the power of classical OT enough to get saltation, but avoid severe collateral damage.
- For reasons to be given, the theory should mildly *discourage* saltation while still permitting it to exist.

**19. Step I:   Abandon classical OT's *minimality restriction* on Faithfulness constraints**

- **Correspondence theory** (McCarthy and Prince 1995), a pillar of classical OT, sensibly uses an **atomistic** approach:
    - ➢ Set up constraints that ban changes in just *one* feature; or deletion of just *one* segment.
- Zuraw (2007):  permit **supraminimal** constraints, but **penalize** their supraminimal status.

**20. *MAP constraints (Zuraw 2007)**

- ***MAP(*x, y*)** assesses a violation to a candidate if a segment belonging to natural class *x* in the input is mapped to a corresponding segment in natural class *y* in the output.

**21. Step II:  a penalty for supraminimality (Zuraw)**

- Appealing to phonetic substance, *MAP constraints are assigned a default ranking as follows:
    - ➢ *MAP constraints banning changes that cover a **larger perceptual distance** are assigned a default ranking higher than constraints banning smaller changes.
- This ranking preference is taken to be a **learning bias** in UG.
    - ➢ Given sufficient evidence in the ambient language, it is possible for learners to subvert the default rankings, but this is harder.

**22. Rankings in \*MAP are based on the P-map**

- Zuraw draws on earlier work by Steriade (2001, 2008) on the **P-map**.
  - ➢ = a compilation of the language learner's phonetic experience concerning the perceptual distance between forms
- It is assumed that language learners consult the P-map in assessing preferred rankings for \*MAP constraints.

**23. Backdrop to \*MAP-cum-P-map theory**

- Essential prediction: alternation prefers to be phonetically minimal; proposed long ago (Vennemann 1972, Kiparsky 1978/1982:65)
- Some literature supporting pressure for phonetically-minimal alternation: Kiparsky (1982), Fleischhacker (2001, 2005), Shademan (2002), Wilson (2006), Zuraw (2007), Lofstedt (2010).
- But as we'll see, we are now in a position for formalize the minimal-alternation-preference, as a bias in phonological learning.

**24. Analysis of Campidanian in \*MAP-cum-P-map theory**

- \*MAP constraint banning [p] ~ [β] alternation: \*MAP(p, β)
- \*MAP constraint banning [b] ~ [β] alternation: \*MAP(b, β)
  - ➢ where p, β stand for the natural classes of voiceless stops and voiced fricatives.

- Key rankings:
  - ➢ \*MAP(b, β) >> \*V[−cont]V          *to keep* /b/ *from spirantizing*
  - ➢ \*V[−cont]V >> \*MAP(p, β)          *to make* /p/ *spirantize*

- By transitivity:
  - ➢ \*MAP(b, β) >> \*MAP(p, β)
  - ➢ … which is **non-P-map-compliant** ([b] is patently closer to [β] than [p] is).
  - ➢ … so such a grammar is a disfavored one, in a sense to be made explicit below

**25. Tableaux for Campidanian**
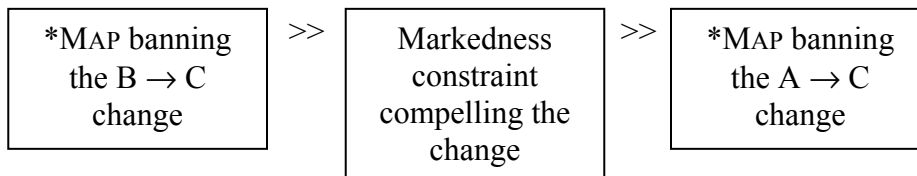
*a.* /p/ *becomes* [β] *intervocalically*

| /apa/ | \*MAP(b, β) | **\*V[−cont]V** | **\*MAP(p, β)** | \*MAP(p, b) |
|---|---|---|---|---|
| ☞ aβa | | | * | |
| *apa | | *! | | |
| *aba | | *! | | * |

*b.* /b/ *is stable*

| /aba/ | *MAP(b, β) | *V[−cont]V | *MAP(p, β) | *MAP(p, b) |
|---|---|---|---|---|
| ☞ aba | | * | | |
| *aβa | *! | | | |

## 26. More generally…

- The known saltations of the form $A \rightarrow C$ *leaping over B* emerge from this ranking:

| *MAP banning the B → C change | >> | Markedness constraint compelling the change | >> | *MAP banning the A → C change |
|---|---|---|---|---|

- … so following *MAP-cum-P-map theory, saltation should be dispreferred.

## V. EVIDENCE THAT SALTATION IS DISPREFERRED

## 27. How could we empirically detect a dispreference for saltation?

- The apparent rarity, and "accidental" origin of saltation might be taken as support.
- We could also check the *learnability* of saltation: do language learners give low credence to saltatory hypotheses when they try to make sense of ambient data?
- In real languages, perhaps the sheer weight of evidence makes saltation learnable.
- But an experimental set-up with an artificial language, things might be different.

## 28. White's experimental work on saltation

- Source: James White (in progress) *Biased learning of phonological alternations*. UCLA dissertation, to be finished ca. May 2013
- Paradigm: artificial-language learning

## 29. Some representative stimuli in White's experiments



[kamap]            [kamavi]
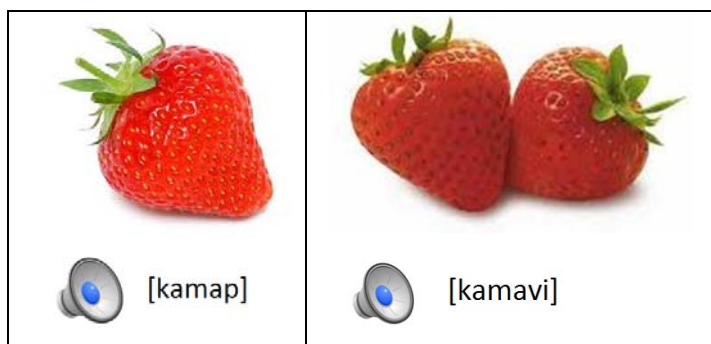
- … where [v] is an English stand-in for Campidanian [β].
- White also used denti-alveolar stimuli; when I say [p b β], assume also [t d ð].

## 30. Experiment 1

- Train subjects with a bunch of forms with [p] ~ [v] alternations.
- Also some filler forms, with invariant stem-final [m, n, l, ɹ, s, ʃ].
- In "arrow diagrams":

```
p             m    n    l    r    s    ʃ
 \
  ↘
   v
```
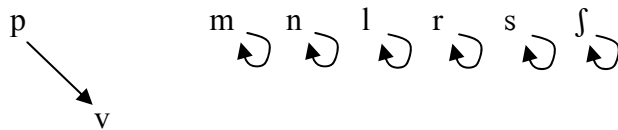
- Check their knowledge:  a two-alternative forced-choice task — give a singular, a changing and a non-changing plural.



| [tinap] | [tinapi] or [tinavi]? |
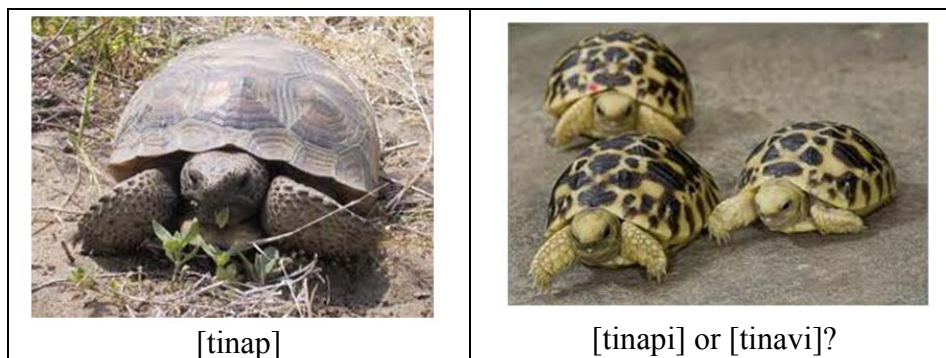
- Use only data from subjects who could score 80% on the trained words.
- Then, ask them *novel words* with [b] and [f] — these have been saltated over.
  - e.g. [tala**b**], [tala___i], [tima**f**], [tima___i]

## 31. White:  Result 1

- Most people get the trained alternation pattern (p → v) pretty accurately.
- They also learn accurately *not* to change the filler consonants [m, n, l, ɹ, s, ʃ].
- The "saltated over" consonants ([b] and [f]):
  - These **often get changed to [v]**, even though subjects were not actually trained to do this.
  - Frequencies of change:

```
          p            b
           \           ⋮
       .96  ↘      .70 ⋮
             ↘         ↓
        f ·········▶ v
            .45
```
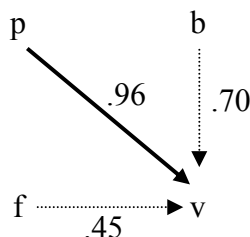
**32. White experiments, part 2: we need to rule out other possibilities**

- The result just given could be simple **generalization**; maybe subjects entertain not just this [p]-specific rule:
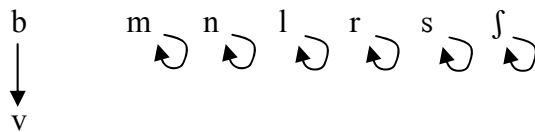
$$\begin{bmatrix} -\text{sonorant} \\ -\text{continuant} \\ -\text{voice} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{voice} \\ +\text{continuant} \end{bmatrix} / V \underline{\quad} V \quad (= p \rightarrow v)$$

  but (to some degree) the more general rule:

$$[-\text{sonorant}] \rightarrow \begin{bmatrix} +\text{voice} \\ +\text{continuant} \end{bmatrix} / V \underline{\quad} V \quad (= \{p, b, f\} \rightarrow v)$$

**33. The crucial control Experiment 2**

- Train instead on **b → v**, and see what happens.



- Saltation is not at stake.
- Generalization is.

**34. Comparative results of Experiments 1 and 2**

- Bold arrow: trained and tested
- Dashed arrows: tested only



*Experiment 1*
*potentially saltatory,*
*massive generalization*

*Experiment 2:*
*not potentially saltatory,*
*modest generalization*
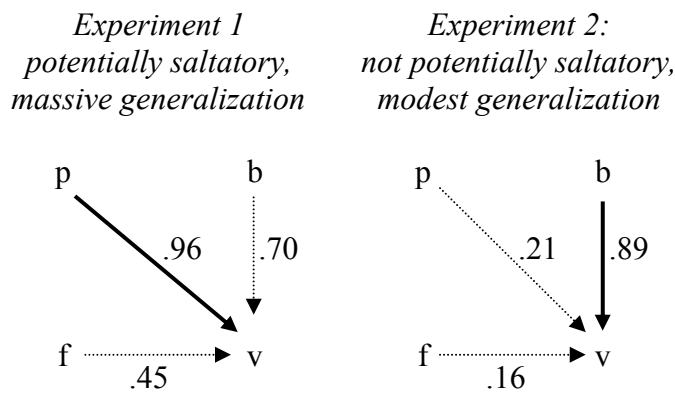
- So both mechanisms seem valid; but the much bigger effect size in Experiment 1 gives us the key result:

***Subjects are reluctant to posit that an alternation is saltatory.***

35. **Even stronger evidence:  Experiment 3**

- White tried even directly training the subjects that /b/ does not change.
    - ➢ Forms like [talab] ~ [talabi].

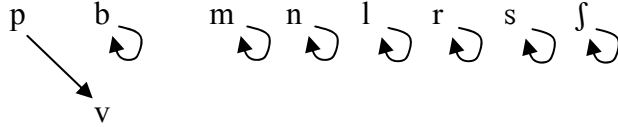p        b⟳        m⟳   n⟳   l⟳   r⟳   s⟳   ʃ⟳

   ↘

        v

- Despite this training, subjects frequently blundered, changing /b/ to [v] in the test forms anyway (about 20% of the time).
- Again, a control experiment again showed that the cause is the saltation, not rule generalization.

VI. IMPLEMENTING AN ANTI-SALTATION BIAS IN PHONOLOGICAL THEORY

36. **What does "bias" *mean*?**

- Plausibly, it means "downgraded for consideration during language acquisition".
- White's experiments suggest that language learners find saltation *a priori* unlikely.
- We want just a bias, not a prohibition:
    - ➢ The saltatory Campidanian pattern *was* learned by children; see Bolognesi (1998:36, 463) for evidence of productivity.
- What theory could give substance to the idea of a learning bias?

37. **A method for implementing bias in learning:  Wilson (2006)**

- Wilson used a learning-bias model to account for the findings of his artificial-language-learning experiment (on palatalization before high vs. mid vowels).
- The model is based on:
    - ➢ a somewhat different constraint-based framework for phonology
    - ➢ results in computational learning theory

38. **Framework:  Maxent grammars**

- First appearance in linguistics:  Goldwater and Johnson (2003)
- Like OT:  GEN (creates candidates) and EVAL (selects winner(s)).
- But EVAL is not based on ranking, but rather a bit of math:
    - ➢ Every constraint has a **weight**; a nonnegative real number.
    - ➢ A formula uses weights and violations to assign a **probability** to every candidate.[10]
    - ➢ This works well for predicting free variation, e.g. of the kind White found.

---

[10] $p(\omega) = \exp(-\Sigma_i \lambda_i \chi_i(\omega))/Z$, where $Z = \Sigma_j \exp(-\Sigma_i \lambda_i \chi_i(\omega_j))$. $p(\omega)$ is predicted probability of candidate $\omega$; $\exp(x)$ is $e$ to the power of $x$; $\Sigma_i$ is summation across all constraints; $\lambda_i$ is weight of the $i$th constraint; $\chi_i(\omega)$ is number of times $\omega$ violates the $i$th constraint; $\Sigma_j$ is summation across candidates.

### 39.  Learning maxent grammars

- The computational learnability situation for maxent (Berger et al. 1996) is excellent:
  - ➢ Given a set of (quantitative) data with candidates, constraints and violations, there is a **unique grammar** (i.e. assignment of weights to constraints) that best fits the data.
  - ➢ … and there are fast algorithms guaranteed to find this grammar.[11]

### 40.  Learning maxent grammars II:  what about soft UG?

- It is feasible to begin the learning process with a **bias** on every constraint.
- A commonly-used bias — the **Gaussian prior** — uses two numbers.
  - ➢ μ:  What is the **preferred weight**?
  - ➢ σ:  How much is the constraint **willing to diverge from** preferred weight?

### 41.  Where will the σ's and μ's come from in this case?

- μ:  from P-map theory
  - ➢ Perceptual closeness of X and Y in an experiment
    - → the μ (preferred weight) of *MAP(X, Y)
- σ:  White set σ at an invariant value of 0.5; no particular basis for letting σ vary.
  - ➢ low enough to let the μ's have a substantial but not overwhelming effect.

### 42.  White's model, step I:  obtaining the μ values for the *MAP constraints

- White used a maxent grammar to model the number of consonant confusions in Wang and Bilger's (1973) perception experiment.

|          |     |     |     |     |     |     |     | Response |     |     |     |     |     |     |     |     |     |
|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Stimulus | p   | t   | k   | b   | d   | g   | f   | θ   | s   | ʃ   | v   | ð   | z   | ʒ   | tʃ  | dʒ  |
| p        | 763 | 23  | 80  | 9   | 3   | 10  | 14  | 6   | 6   | 4   | 4   | 2   | 3   | 5   | 1   | 4   |
| t        | 25  | 795 | 31  | 1   | 9   | 6   | 4   | 4   | 12  | 3   | 0   | 2   | 2   | 2   | 22  | 9   |
| k        | 66  | 36  | 753 | 2   | 7   | 18  | 7   | 5   | 4   | 10  | 3   | 1   | 3   | 5   | 7   | 6   |
| b        | 55  | 12  | 28  | 608 | 19  | 87  | 46  | 9   | 9   | 2   | 43  | 6   | 3   | 3   | 2   | 4   |
| d        | 9   | 16  | 7   | 10  | 763 | 13  | 8   | 13  | 12  | 2   | 15  | 33  | 14  | 9   | 3   | 9   |
| g        | 16  | 10  | 44  | 22  | 7   | 762 | 13  | 5   | 3   | 4   | 20  | 4   | 3   | 4   | 2   | 12  |
| f        | 56  | 0   | 11  | 22  | 9   | 13  | 712 | 61  | 10  | 3   | 25  | 8   | 0   | 1   | 0   | 2   |
| θ        | 11  | 14  | 9   | 9   | 42  | 10  | 73  | 656 | 28  | 7   | 23  | 36  | 13  | 1   | 1   | 3   |
| s        | 7   | 15  | 12  | 1   | 3   | 4   | 15  | 23  | 828 | 14  | 1   | 1   | 11  | 0   | 1   | 0   |
| ʃ        | 9   | 17  | 10  | 7   | 9   | 6   | 6   | 1   | 18  | 788 | 4   | 2   | 4   | 18  | 19  | 18  |
| v        | 5   | 5   | 8   | 38  | 22  | 19  | 48  | 7   | 11  | 4   | 637 | 104 | 12  | 7   | 1   | 5   |
| ð        | 8   | 7   | 5   | 38  | 77  | 4   | 10  | 37  | 6   | 7   | 136 | 556 | 34  | 8   | 1   | 2   |
| z        | 4   | 8   | 15  | 4   | 12  | 10  | 8   | 22  | 42  | 8   | 54  | 35  | 687 | 17  | 4   | 3   |
| ʒ        | 9   | 3   | 9   | 1   | 13  | 7   | 4   | 7   | 15  | 30  | 24  | 13  | 22  | 752 | 4   | 20  |
| tʃ       | 10  | 17  | 14  | 5   | 2   | 7   | 2   | 1   | 5   | 14  | 2   | 1   | 0   | 5   | 727 | 121 |
| dʒ       | 16  | 20  | 31  | 2   | 25  | 17  | 4   | 1   | 11  | 10  | 3   | 2   | 8   | 7   | 86  | 689 |

---

[11] In our work we used the Maxent Grammar Tool (www.linguistics.ucla.edu/people/hayes/ MaxentGrammarTool/, a user-friendly version of Wilson's software) to find the weights.

- This gave him empirically-grounded μ values, as follows.

| Constraint | μ | Constraint | μ |
|---|---|---|---|
| *MAP(p, f) | 1.61 | *MAP(b, f) | 1.32 |
| *MAP(p, b) | 1.98 | *MAP(f, v) | 2.45 |
| *MAP(p, v) | 3.57 | *MAP([+son], [−son]) | 14.63 |
| *MAP(b, v) | 0 | | |

- The markedness constraints were assigned μ of zero, by fiat.

**43. Step 2: train the model as if it were an experimental subject**

- Constraints, σ, μ: as above
- Training data = exactly what the subjects got. Separate simulation for each experiment.

**44. How the weights changed during learning**

- Training data was Expt. 3: [p] becomes [v], [b] stays [b]; filler forms stay unchanged.
- Weights changed like this:

| Constraint | μ | Post-learning weight |
|---|---|---|
| *V[−voice]V | 0 | 2.04 |
| *V[−contin]V | 0 | .48 |
| *MAP(p, f) | 1.61 | 1.95 |
| *MAP(p, b) | 1.98 | 2.74 |
| **\*MAP(p, v)** | **3.57** ⟶ | **1.04** |
| **\*MAP(b, v)** | **0** ⟶ | **1.51** |
| *MAP(b, f) | 1.32 | 1.51 |
| *MAP(f, v) | 2.45 | 2.45 |
| *MAP([+son], [−son]) | 14.63 | 14.63 |

**45. The effect of bias in learning**

- We can compare the result with what happens if we run the learning model *without* bias:
- Without bias, *MAP(p, v) gets zero weight; it is violated extensively.
  ➢ In the bias simulation, its μ value (3.57) kept it fairly high (1.04).
- Without bias, *MAP(b, v) gets a very high weight — it's needed to retain [b] as [b].
  ➢ In the bias simulation, its μ value (0) kept it fairly low (1.51).

**46. Step 3: test the model as if it were an experimental subject**

- I.e. compare the frequency with which the subjects made a particular guess with the frequency with which the model makes it.
- In general, White's model matches the data rather closely.

**47. Sample result I**

- Train [p] → [v], [b] → [b]  (Expt. 3)
  - ➢ Model, like people, guesses about 20% that b → v.
  - ➢ Reason:  *MAP(b, v), with a prior weight of 0, fails to rise high enough during learning to block this change.

**48. Sample result II**

- In Expt. 1 ((31)), subjects offered [f] → [v] (.45) less often than [b] → [v] (.7) — why?
  - ➢ Wang and Bilger's data show that [b] is closer to [v] than [f] is to [v].
  - ➢ So, calculated µ values were:  *MAP(f, v), 2.45; *MAP(b, v), 0
  - ➢ So, in learning, *MAP(f, v) stays high-ish
  - ➢ So model predicts that [f] is less likely to alternate.

**49. White's model fit (all experiments, all predictions vs. observed)**
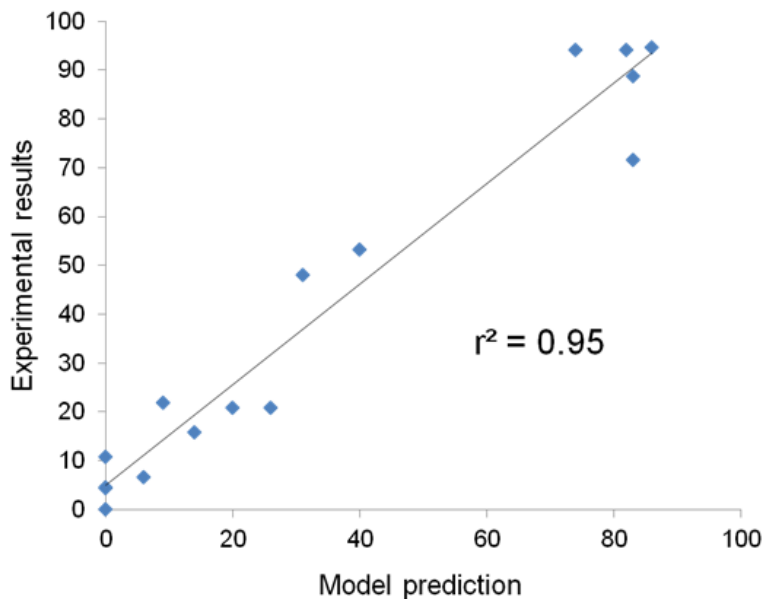


VII. CONCLUDING GENERAL REMARKS

**50. The prospects for soft UG**

- Soft UG has always seemed intuitively sensible to many people but has been frustratingly fuzzy; how to make it a rigorous theory?
- A possible answer is:
  - ➢ Soft UG is embodied as a learning bias.
  - ➢ Maxent provides the right formal language.[12]

---

[12] This is a bit parochial; the real hero is probability theory, from which the maxent principles are deduced.

➢ Experiments provide data about learning difficulty that are hard to obtain "in the wild".

## 51.  Goals of phonological theory

- Permit analysis of intricate systems — especially when evidence tells us such systems are productively internalized.
- Also:  say what is **hard** for people — where is acquisition liable to fail?
  - ➢ The errors of White's subjects
  - ➢ The drastic grammar change that created Campidanian saltation (see (6)-(7))
- This is not a new research goal — Kiparsky (1970/1982: 59-60).
  - ➢ Kiparsky's pioneering research on opacity was focused precisely on the *breakdowns* that diagnose acquisition difficulty.

## 52.  The Kiparskyan criterion and current research on opacity

- Opinion:  in Harmonic-Serialism-cum-candidate-chain theory John McCarthy and colleagues (see e.g. McCarthy 2007, 2008, 2011) have created **the best theory of phonological opacity that has ever existed**.
- *Every* classical opaque system, e.g. as embodied in standard problem sets, has a strikingly elegant analysis in this theory.
  - ➢ This analysis is, normally, a recapitulation of the system's diachronic origins.[13]
- But, in the development of this theory, has Kiparsky's original motivation for investigating opacity been forgotten?

## 53.  Opacity is often diachronically unstable

- Many classical problem sets are in real life riddled with exceptions — evidence, I think, that the system may have restructured.
- Examples, with work discussing the possibility of restructuring:
  - ➢ Catalan cluster simplification (Wheeler 2005)
  - ➢ Lardil apocope/consonant deletion/raising (Richards 1997, Rounds 2011)
  - ➢ Odawa stress/syncope (Bowers 2012)
  - ➢ Polish raising and devoicing (Sanders 2003)
  - ➢ Serbo-Croatian /a/ epenthesis and /l/ vocalization (Bochner 1981)
  - ➢ Ukrainian raising/palatalization (Kenstowicz and Kisseberth 1977, 161-2)
  - ➢ Yawelmani vowel harmony (Hansson and Sprouse, ms.)
  - ➢ Yidiɲ lengthening and apocope (Hayes 1999)
- Such cases may be problematic for a theory in which opacity flows as smooth as oil — sometimes a theory can be *too* good!
- I think we should strive for a theory of opacity that predicts:
  - ➢ when opacity is stable[14]
  - ➢ when it is vulnerable to breakdown

---

[13] Indeed, the derivations sometimes recapitulate history down to the very details of reduction and spreading that phonetically underlie sound change — see McCarthy (2008, 2011).

[14] I think sometimes it *is* stable, e.g. in single-segment chain shifts.

- The key to such a (future) theory, I think, is likely to be the global similarity relations among the overt forms in the paradigm, not the local similarity relations that governed how the pattern was historically created.

## 54.  Summing up

- Saltation is a minor topic in phonology in terms of its empirical scope.
- Yet it is perhaps informative concerning the usefulness of theoretical approaches to phonology.
- Here, the concepts that seem (to me) to be winners are:
  - ➢ Appeal to phonetic substance (here, *Map-cum-P-map) for explanation in phonology
  - ➢ Gradient versions of UG, expressed as learning biases
  - ➢ "Fuzzy" theories made rigorous through the use of appropriate mathematics

## References

Baković, Eric. 2007.  A revised typology of opaque generalizations. *Phonology 24*:217-259.

Berger, Adam L., Stephen A. Della Pietra, and Vincent J. Della Pietra. 1996. A maximum entropy approach to natural language processing. *Computational Linguistics* 22:39–71.

Bochner, Harry. 1981. The l → o rule in Serbo-Croatian.  In G. N. Clements, ed., *Harvard Studies in Phonology*, Vol. 2.

Bolognesi, Roberto. 1998. *The Phonology of Campidanian Sardinian*.  The Hague:  Holland Academic Graphics.

Bowers, Dustin. 2012) Phonological restructuring in Odawa.  MA thesis, UCLA. https://sites.google.com/site/dustinbowerslinguist/papers

Bradshaw, Mary. 1999. Unrecoverable origins. In Hermans and van Oostendorp, pp. 51-80.

Fleischhacker, Heidi. 2001. Cluster-dependent epenthesis asymmetries. *UCLA Working Papers in Linguistics* (*Papers in phonology* 5) 7.71–116.

Fleischhacker, Heidi. 2005. Similarity in phonology: Evidence from reduplication and loan adaptation. Los Angeles: UCLA dissertation.

Goldwater, Sharon, and Mark Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In Proceedings of the Stockholm Workshop on Variation within Optimality Theory, ed. by Jennifer Spenader, Anders Eriksson, and Osten Dahl, 111–120.

Hansson, Gunnar and Ronald Sprouse. Ms. Factors of change: Yowlumne vowel harmony then and now.

Hermans, Ben and Marc van Oostendorp, eds. 1999. *The derivational residue in phonological Optimality Theory*. Amsterdam:  John Benjamins.

Hayes, Bruce. 1999. Phonological restructuring in Yidiɲ and its theoretical consequences. In Hermans and van Oostendorp, pp. 175-205.

Ito, Junko and Armin Mester. 1997. Correspondence and compositionality: The Ga-gyō variation in Japanese phonology.  In Roca, Iggy, ed. Derivations and Constraints in Phonology. Oxford University Press.

Ito, Junko and Armin Mester. 2003. On the sources of opacity in OT: coda processes in German. In Caroline Féry and Ruben van de Vijver (eds.), *The Syllable in Optimality Theory*, Cambridge University Press. 271-303.

Kenstowicz, Michael and Charles Kisseberth. 1977. *Topics in Phonological Theory*.  New York: Academic Press.

Kiparsky, Paul. 1970. Historical linguistics. in J. Lyons (ed.) *New horizons in linguistics.* Harmondsworth:  Penguin.

Kiparsky, Paul. 1978. Historical linguistics.  Reprinted as chapter 4 of Kiparsky 1982.

Kiparsky, Paul. 1982. *Explanation in phonology*. Dordrecht:  Foris.

Lass, Roger. 1997. *Historical linguistics and language change.*  Cambridge:  Cambridge University Press.

Lofstedt, Ingvar. 2010. Phonetic effects in Swedish phonology: Allomorphy and paradigms. Los Angeles: UCLA dissertation.

Lubowicz, Anna. 2002. Derived Environment Effects in Optimality Theory. *Lingua* 112, 243-280.

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

McCarthy, John. 2008. The gradual path to cluster simplification. *Phonology* 25, 271-319.

McCarthy, John. 2011. Perceptually grounded faithfulness in Harmonic Serialism. *Linguistic Inquiry* 42, 171-183.

McCarthy, John  and Alan Prince. 1995. Faithfulness and Reduplicative Identity, in *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. Ed. by Jill Beckman, Suzanne Urbanczyk and Laura Walsh Dickey. Pp. 249–384.

Minkova, Donka. 1993. On leapfrogging in historical phonology.  In Jaap Van Marle, ed.,  *Historical Linguistics 1991: Papers from the 10th International Conference on Historical Linguistics, Amsterdam, 12-16 August 1991*.  Amsterdam:  John Benjamins.

Pilch, Herbert. 1966. Das Lautsystem der hochdeutschen Umgangssprache. *Zeitschrift für Mundartforschung* P

Prince, Alan and Paul Smolensky 1993/2004.  Optimality theory: Constraint interaction in generative grammar.  .  Cambridge, MA: Blackwell. [Technical report CU-CS-696–93, Department of Computer Science, University of Colorado at Boulder, and Technical report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, 1993.]

Richards, Norvin. 1997. Old and New Lardil. *MIT Occasional Papers in Linguistics* 13.

Robinson, Orrin W. 2000. *Whose German? The* ach*/*ich *alternation and related phenomena in 'standard' and 'colloquial'*.  Amsterdam:  John Benjamins.

Rounds, Erich. 2011. Word final phonology in Lardil: Implications of an expanded data set. *Australian Journal of Linguistics* 31:327-350.

Rubach, Jerzy. 1993. *The lexical phonology of Slovak*. New York: Oxford University Press.

Sanders, Nathan. 2003) Opacity and sound change in the Polish lexicon.  UC Santa Cruz dissertation.

Schuh, Russell. Ms. The history of labial and velar obstruents in Kanuri. Ms., Dept. of Linguistics, UCLA, Los Angeles, CA.  http://www.linguistics.ucla.edu/people/schuh/Papers/B36_kanuri_lab_vel_history.pdf

Shademan, Shabnam. 2002. Epenthetic vowel harmony in Farsi. Los Angeles: University of California, Los Angeles masters thesis.

Smolensky, Paul. 1995. On the structure of the constraint component Con of UG. Paper presented at the University of California, Los Angeles. Handout available as ROA-86 from the Rutgers Optimality Archive.

Steriade, Donca. 2001. Directional asymmetries in place assimilation: A perceptual account. In Hume, Elizabeth, and Keith Johnson (eds.) 2001. *The role of speech perception in phonology*. San Diego, CA: Academic Press.

Steriade, Donca. 2008. The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In *The nature of the word: Studies in honor of Paul Kiparsky*, ed. by Kristin Hanson and Sharon Inkelas, 151–80. Cambridge, MA: MIT Press.

Vennemann, Theo. 1972. Phonetic analogy and conceptual analogy. In Vennemann, Theo, and Terence H. Wilbur (Eds.), *Schuchardt, the Neogrammarians, and the Transformational Theory of phonological change: Four essays*. Frankfort, Athenaum. p. 181-204.

Wang, Marilyn D. and Robert C. Bilger. 1973. Consonant confusions in noise: a study of perceptual features. *Journal of the Acoustical Society of America* 54:1248-1266.

Wheeler, Max W. 2005. *The phonology of Catalan*. Oxford: Oxford University Press.

White, James. 2012. On the learnability of saltatory alternations.  Poster given that the 13th Conference on Laboratory Phonology, Stuttgart. www.linguistics.ucla.edu/people/grads/jwhite/conferences/White_LabPhon13_poster.pdf

White, James. In progress. Biased learning of phonological alternations. Ph.D. dissertation, UCLA.

Wilson, Colin. 2006. Learning phonology with substantive bias: an experimental and computational investigation of velar palatalization. *Cognitive Science* 30:945–982

Zhirmunskii, Viktor Maksimovich. 1962. *Deutsche Mundartkunde:  Vergleichenda Laut- und Formenlehre de deutschen Mundarten*. Berlin: Akademie-Verlag.

Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog. *Language* 83: 277-316.