# A New Approach to the Origin of Germanic Strong Preterites Implications for the Phonology Morphology Interface and Diachronic Learnability

Ryan Sandell & Sam Zukoff
University of California, Los Angeles & Massachusetts Institute of Technology
ryan.sandell@gmail.com, szukoff@mit.edu

1 November 2014

## 1. Overview of the Problem and Questions

1. 1. How is the system of Germanic strong verbs, as best represented by Gothic, to be analyzed synchronically (sections 2–3)?
    - This system marks past tense through vowel change (ablaut), and less commonly, reduplication.
    - We argue that the attested patterns of past tense marking result from the relative rankings of faithfulness constraints with respect to two morphological constraints, Realize-Morpheme (Kurisu 2001) and Anti-Ident (Crosswhite 1999).
1. 2. How is the ancestor of that system in Proto-Germanic, inherited from Proto-Indo-European, to be analyzed synchronically (section 4)?
    - This system consistently marks past tense with reduplication (thus a /Red/ morpheme), in which vowel alternations depend upon accentual properties.
1. 3. How can the changes between the inherited and attested systems be motivated and modeled?
    - This is a crux of Germanic historical linguistics: how can a grammar with consistent reduplication become a grammar without it?
    - There is a "gap" between the grammars of earlier and later Proto-Germanic that needs to be filled in — how did one become the other?
    - We propose that the change in prosodic system, from a lexical accent system to a purely phonological accent system, both made reduplication difficult to learn and allowed for previously phonologically driven vowel alternation patterns to serve a morphological function.
1. 4. How are tools of OT grammar learning properly used, and their results properly interpreted, when trying to capture diachronic changes (section 5)?
    - How do Maximum Entropy models behave when provided with a constraint analysis fitted to selecting one winner, but a set of "adult" observed forms fit to a different analysis?
    - Either effects of **learning bias** (Hayes et al 2009b) or **acquisition order** (Boersma and Levelt 2000) may be sufficient to capture the desired diachronic trajectory.

## 2. Structure and Patterns of the Strong Verb System: The Gothic Evidence

### 2. 1. Basics of Gothic "Strong" Verbal Inflection (consult generally Lambdin 2006 or Bennett 1965 [1980])

2. 1. 1. Gothic verbs divide into two large classes: STRONG and WEAK verbs, which are largely phonologically distinguishable on the basis of their present tense stem.

2. 1. 2. Both "strong" and "weak" verbs form different stems for present and past tense.

2. 1. 2. 1. Weak verbs derive their preterites with the "dental" suffix, which is /-d-/ in the singular and /-deːd-/ in the plural.

- This weak verb formation corresponds to English verbs that form their preterites with the suffix /-d/.

2. 1. 2. 2. Strong verbs derive their preterites through various vowel changes and reduplication (see further below).

- These formations correspond to English "irregular" preterites.

2. 1. 2. 3. Gothic verbs obligatorily take suffixes marking person and number (sg. and pl., marginal dual); these are identical for strong and weak verbs in the present, and in the preterite plural, but there are different endings for the respective classes in the pret.sg.

(1) Gothic Present Tense Inflectional Endings

| | Sg. | Pl. |
|---|---|---|
| 1 | /-a/ | /-am/ |
| 2 | /-i(ː)s/ | /-eːθ/ |
| 3 | /-i(ː)/ | /-and/ |

(2) Gothic Past Tense Inflectional Endings

| | Sg. | Pl. |
|---|---|---|
| 1 | /-a/ / -Ø | /-um/ |
| 2 | /-eːs/ / /-t/ | /-uθ/ |
| 3 | /-a/ / -Ø | /-un/ |

2. 1. 3. The system of strong verbs is further divided into seven classes, distinguishable by phonological properties of the verbal root. These properties in turn determine how the preterite is formed.

(3)

| General Root Shape | Class | PRESENT | 3.SG.PRET. | 1.PL.PRET. | GLOSS |
|---|---|---|---|---|---|
| /CeRC/ | I | /bejt-an/ [bɛjtan] | /bajt/ [bajt] | /bjt-um/ [bɪtʊm] | 'bite' |
| | II | /kews-an/ [kɪwsan] | /kaws/ [kaws] | /kws-um/ [kʊsʊm] | 'choose' |
| | III | /bend-an/ [bɪndan] | /band/ [band] | /bnd-um/ [bʊndʊm] | 'bind' |
| /CeC/ | IV | /nem-an/ [nɪman] | /nam/ [nam] | /neːmum/ [neːmʊm] | 'take' |
| | V | /geb-an/ [gɪβan] | /gab/ [gaf] | /geːbum/ [geːfʊm] | 'give' |
| /CaC/ | VI | /dab-an/ [daβan] | /daːb/ [doːp] | /daːb-um/ [doːβʊm] | 'happen' |
| /CVːC/ | VIIa | /hald-an/ [haldan] | /he-hald/ [hɛhald] | /he-hald-um/ [hɛhaldʊm] | 'hold' |
| | VIIb | /leːt-an/ [leːtan] | /le-laːt/ [lɛloːt] | /le-laːt-um/ [lɛloːtʊm] | 'let' |
| | VIIc | /flaːk-an/ [floːkan] | /fe-flaːk/ [fɛfloːk] | /fe-flaːk-um/ [fɛfloːkʊm] | 'bewail' |

2. 1. 1. Note that the root vocalism in the plural of Classes I-III represents the vocalization of an underlying sonorant:

2. 1. 1. 1. Class I = roots with underlying medial /j/: [bɪtum] ← /bjt-um/

2. 1. 1. 2. Class II = roots with underlying medial /w/: [kʊsum] ← [kws-um]

2. 1. 1. 3. Class III = roots with underlying medial non-glide sonorant /m,n,r,l/: [bʊndum] ← /bnd-um/

2. 1. 2. The most prominent division among root classes is that between roots showing a short front vowel ([ɪ] or [ɛ]) in the present (Classes I–V), versus those showing any other vowel, most commonly [a], but also [eː] and [oː] (Classes VI and VII).

2. 1. 3. Clasess I–V, the first of these two groupings, all form their pret.sg. stem by vowel change to [a].

2. 1. 3. 1. Roots ending in a sequence $C_{[+son]}C$ (Classes I–III) show vowel deletion in the pret.pl.

- The division between I–III is whether the sonorant is /j/, /w/, or a liquid or nasal.

2. 1. 3. 2. Roots ending in a single consonant show vowel lengthening to [eː] in the pret.pl.

- The division between IV and V is whether final consonant is a sonorant (Class IV) or an obstruent (Class V).

2. 1. 4. Classes VI and VII show no difference between the stem of the pret.sg. and the pret.pl.

2. 1. 4. 1. Class VI shows vowel lengthening to [oː] in the pret.sg. and pret.pl. Roots belonging to Class VI (crucially) end in a single consonant.

2. 1. 4. 2. Class VII shows reduplication in the pret.sg. and pret.pl. Class VII roots are "heavy": they terminate in two consonants ([hald-]) or a long vowel and a consonant ([leːt-]).

## 3. Phonological Repairs to Satisfy Realize-Morpheme and Anti-Ident: The Gothic Analysis

3. 1. What triggers or drives the patterns of vowel change seen in the root between the present and the preterite? Why does reduplication occur at all?

3. 2. The patterns clearly cannot be triggered by inflectional endings themselves (like the German pl. marker /-e/, e.g., *Fuß* : pl. *Füße*)

- There are no such changes in weak verbs that share the same inflectional endings in the pret.pl. as strong verbs!
- Why should the phonological realizations of some feature from an ending be realized in such diverse ways?

3. 3. What seems to be needed is a paradigmatic base : derivative relationship, and a tight connection between phonology and morphology. These criteria can be implemented in several ways.

3. 3. 1. Morphological rules sensitive to phonological context (Albright 2002, Albright and Hayes 2003).

3. 3. 2. Morphological schemas that reference phonological information in their components (Booij 2010).

3. 3. 3. Phonological repairs triggered by high-ranking morphological constraints.

3. 4. In general conception, these choices are theoretically similar: they all require a network of relations between morphosyntactic forms that packages phonology and morphology together. Since learnability questions are, at present, most easily addressed with constraint grammars, we adopt the third approach.

3. 5. Relevant morphological contrasts are the following:

3. 5. 1. The preterite stem must be phonologically distinct from the present stem. This principle is often noted for Germanic strong verbs (e.g., Meid 1971), and always holds.

3. 5. 2. The stems pret.sg. and pret.pl. should be phonologically distinct (although this is not always borne out, as in Class VI and VII).

3. 5. 3. We assume that these contrasts must hold at the stem level.

3. 5. 4. We employ existentially quantified faithfulness constraints, rather than universally quantified faithfulness constraints (Struijke 2002).

3. 6. The first requisite (preterite stems are always different from the present stem) can be captured through a constraint requiring a difference in morphosyntactic category to be phonologically expressed. Kurisu (2001: 39) defines a constraint Realize-Morpheme (RM) as follows:

(4) Realize-Morpheme (RM):
Let α be a morphological form, β be a morphosyntactic category, and F(α) be the phonological form from which F(α+β) is derived to express a morphosyntactic category β. Then RM is satisfied with respect to β iff F(α+β) ≠ F(α) phonologically.

(5) Class I Pret.Sg. (applies also to Class II-V pret.sg.)

| /bɛjt, Pret.Sg/ Pres.: [bɛjt-] | | RM | ∃-Max-V | ∃-Ident[-back] |
|---|---|---|---|---|
| a. | bɛjt | *! | | |
| b. | bi̯t | | *! | * |
| c. ☞ | bajt | | | * |

3. 7. For the second requirement, what is needed is a constraint that evaluates whether two given morphological outputs from a given base are formally identical or not. This can be enforced through a constraint requiring non-identity of two forms. Effects of paradigmatic morphological contrast are amply reported (Crosswhite 1999, Crosswhite 2001 Kenstowicz 2002, Ito and Mester 2004). Crosswhite (2001: 155) defines the constraint Anti-Ident as follows:

(6) Anti-Ident: For two forms, $S_1$ and $S_2$, where $/S_1/ \neq /S_2/$, $\exists\, \alpha, \alpha \in [S_1]$, such that $\alpha \neq \mathscr{R}(\alpha)$, where $\mathscr{R}$ is a correspondence relation between two strings.

3. 7. 1. Anti-Ident thus evaluates whether or not complete segmental identity holds between one output form and another output form. There should be some segment α, which is a member of $S_1$ that is not identical to its correspondent in $S_2$.

3. 7. 2. For example, Crosswhite (1999: 10) compares (Trigrad) Bulgarian singular forms with their plurals; underlying /zorno/ surfaces as *zorna* 'grain, pl.', satisfying constraints on vowel reduction, but surfaces as *zorno* 'grain, sg.' to satisfy higher-ranked Anti-Ident, but violates phonological constraints on vowel reduction.

3. 7. 3. In analyzing the Gothic system, we will evaluate the pret.pl. stem with respect to the pret.sg. stem

(7) Class I Pret. Pl. (applies also to Class II-III pret.pl.)

| /bɛjt, PRET., um/ PRES.: [bɛjt-], Sg. : [bajt-] | | RM | ANTI-IDENT | ∃-MAX-V | ∃-IDENT[-back] |
|---|---|---|---|---|---|
| a. | bɛjtum | *! | | | * |
| b. ☞ | bj̩tum | | | * | * |
| c. | bajtum | | *! | | * |

3. 8. Lengthening (penalized by ∃-IDENT[-long]) and reduplication (penalized by ∃-INTEGRITY) are evidently less optimal repairs than backing (penalized by ∃-IDENT[-back]) and vowel deletion (penalized by ∃-MAX-V), because they occur only when the preferred repairs are blocked by higher-ranking morphological or phonotactic constraints.

3. 8. 1. Lengthening (i.e. ∃-IDENT[-long] violation) applies:

1. in Class IV-V pl., because vowel deletion (violation of ∃-MAX-V) would create new complex onset sequences (e.g., [gb-], [grb-]), which we penalize with a *COMPLEX.
2. Another peculiarity of the Gothic patterns: in cases where the root vowel of the present stem is [-back], it may become [+back] in the preterite, apparently satisfying RM (e.g. pres. *beitan* : pret. *bait*), but [+back] vowels in the present never become [-back] (pres. [graban] : pret. [gro:f], not X[grɪf]).
3. in Class VI sg. & pl., because high-ranking ∃-IDENT[+back] enforces retention of the [+back] feature.

(8) Class V Pret.Pl.

| /geb, PRET., um/ PRES.: [gɪb-], Sg. : [gab-] | | RM | *COMPLEX | ANTI-IDENT | ∃-IDENT[-long] | ∃-MAX-V | ∃-IDENT[-back] |
|---|---|---|---|---|---|---|---|
| a. | gɪbum | *! | | | | | |
| b. | gbum | | *! | | | * | * |
| c. | gabum | | | *! | | | * |
| d. ☞ | ge:bum | | | | * | | |

(9) Class VI Pret.Sg.

| /dab, PRET./ PRES.: [dab-], | | RM | *COMPLEX | ∃-IDENT[+back] | ANTI-IDENT | ∃-IDENT[–long] | ∃-MAX-V | ∃-IDENT[-back] |
|---|---|---|---|---|---|---|---|---|
| a. | dab | *! | | | | | | |
| b. | db | | *! | | | | * | |
| c. | dɛb | | | *! | | | | |
| d. ☞ | do:b | | | | | * | | |

3\. 8. 2. In order to satisfy REALIZE-MORPHEME, Gothic shows violations of the following faithfulness constraints:

3\. 8. 2. 1. ∃-IDENT[-back] (pret.sg. of Classes I-V)

3\. 8. 2. 2. ∃-MAX-V (pret.pl. of Classes I-III)

3\. 8. 2. 3. ∃-IDENT[-long] (pret.pl. of Classes IV-VI, pret.sg. of Class VI)

3\. 8. 2. 4. ∃-INTEGRITY (Class VII)

3\. 8. 3. Violations of ∃-INTEGRITY are only tolerated in Class VII, thereby creating reduplicated forms, because lengthening of the root vowel would create superheavy syllables, which are banned by high-ranking ⋆SUPERHEAVY.

3\. 8. 4. An ∃-INTEGRITY violation is thus the last-resort repair.

(10) Class VII Pret.Pl.

| /hajt, PRET., um/ PRES.: [hajt-], Sg.: [hɛhajt-] | | | RM | ⋆SPRHVY | IDENT[+back] | ∃-INTEGRITY | ANTI-IDENT | ∃-IDENT[-long] | ∃-MAX-V | ∃-IDENT[-back] |
|---|---|---|---|---|---|---|---|---|---|---|
| a. | | hajtum | ∗! | | | | | | | |
| b. | | hjtum | | | ∗! | | | | ∗ | |
| c. | | hɛjtum | | | ∗! | | | | | |
| d. | | ho:j.tum | | ∗! | | | | ∗ | | |
| e. | ☞ | hɛhajtum | | | | ∗ | ∗ | | | |

3\. 9. **General ranking**: REALIZE-MORPHEME, ⋆SUPERHEAVY, ⋆COMPLEX, ∃-IDENT[+back] ≫ ∃-INTEGRITY-IO ≫ ANTI-IDENT ≫ ∃-IDENT(-long)-IO, ≫ ∃-MAX-V-IO ≫ ∃-IDENT[-back]-IO

3\. 10. Summary: alternations between the present stem and the preterite sg. and pl. stems in Gothic strong verbs are explicable by the ranking of phonological constraints with respect to two morphological constraints, REALIZE-MORPHEME and ANTI-IDENT. Weak verbs select for the preterite morpheme /-(de:)d/ and thereby trivially satisfy REALIZE-MORPHEME.

## 4. The Ancestor of the Gothic System: Evidence from Greek and Sanskrit

4\. 1. Three facts about the Indo-European "perfect" (resultative/anterior aspect):

4\. 1. 1. The alternation between [a] and Øseen in forms like [ˈbajt] ∼ [ˈbit] must continue a Proto-Indo-European *[o] ∼ Ø.

4\. 1. 1. 1. Gk. 3.sg.perf. [lelóipe] 'remains' : 3.pl.perf. [lélipon] < PIE *[lelójpe] : *[lelipŕ̥]

4\. 1. 1. 2. Skt. [cəká:rə] 'has made' : [cəkrúr] < PIE *[$k^w$e$k^w$óre] : [$k^w$e$k^w$ŕ̥]

4\. 1. 2. The "perfect" in Greek and Sanskrit has a CV- reduplicant. In Gothic, reduplication appears only in Class VIII, and the root shows no vowel alternation!

4\. 1. 3. The Indo-European *[o] and Ø vowels are derivable from underlying */e/ according to accent patterns.

4\. 1. 3. 1. /e/ → Ø/ _$C_0$V́

4\. 1. 3. 2. /e/ → [ó] / e$C_0$_$C_0$e

4. 2. A few Sanskrit perfects with "missing" reduplication are precisely paralleled in Gothic.

4. 2. 1. Sanskrit perf.pl. of the form [$C_1$eː$C_2$-]

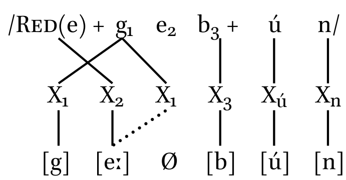
1. 3.sg. [səsáːdə] : 3.pl. [seːdúr] (/səd/ 'sit')
2. 3.sg. [nənáːmə] : 3.pl. [neːmúr] (/nəm/ 'bow')
3. 3.sg. [pəpáːcə] : 3.pl. [peːcúr] (/pəc/ 'cook')

4. 2. 2. Indeed, Skt. [neːmúr] and [seːdúr] are matched by Goth. [ˈneːmun] and [ˈseːtun].

- NB: NOT directly continuing PIE *[neːmúr] and *[seːdúr] (which would rather give Skt. $^X$[naːmúr] and $^X$[saːdúr]), but synchronically built as such un the separate languages, having similar constraint rankings.

4. 2. 3. Because the vowel deletion patterns are transparent, the consonant deletion and compensatory lengthening allow for recovery of underlying reduplication.

(11) /Red-CeC-/ → [CeːC-] Compensatory Lengthening




4. 3. We can therefore assume that Germanic inherited PIE 3.sg.perf. *[sesóde] : 3.pl.perf. [seːdúr]. But how did PIE *[sesóde] turn into Gothic [ˈsad]?

4. 4. There is no surface-oriented phonological change ("sound change") that can explain the loss of the reduplication.

4. 5. There is one important difference between Sanskrit and Germanic, however: Germanic has developed a fixed initial stress accent, replacing the Indo-European mobile accent.

4. 5. 1. Sanskrit 3.sg.perf. [səsáːdə] : 3.pl. [seːdúr] is transparent: accentuation feeds vowel deletion which feeds consonant deletion and compensatory lengthening.

4. 5. 2. A Proto-Germanic 3.sg.pret. *[ˈsesate] : 3.pl. *[ˈseːtun] is opaque: the accentual conditions for vowel deletion are gone!

4. 6. At some point, Proto-Germanic learners faced the paradox of [ˈsesate] : [ˈseːtun]. What could they do?

## 5. Learnability Problems: How does the system of 4. become the system of 2–3?

5. 1. Our learning objective: find conditions under which the *observed* winners *do not* win under a grammar.

5. 1. 1. Is there some set of learning conditions under which observed data will lead learners to choose a different output?

5. 1. 2. A question: do off-line and on-line learning models make substantially different predictions?

5. 2. Our method: pair the constraint grammar developed for Gothic in **3.** with the reconstructed observed forms of **5.**

5\. 2. 1. The basic issue all the forms of **4.** (*[ˈbebidun], *[ˈne:mun])are structurally ambiguous: do they contain an underlying /Red/ morpheme (preserved by high-ranking Max-BR), or do they reflect a phonological Integrity violation?

5\. 2. 1. 1. At present, we set aside the problem of how the appropriate URs and constraints themselves could be induced from surface forms; we focus only on the learning of constraint rankings/weighting.

5\. 2. 1. 2. We assume that learners are arriving to URs without a /Red/ morpheme, and attribute any surface reduplication to Integrity violations.

5\. 2. 2. Here is a tableaux of violations for /bejd, Pret, un/, with the "adult" form indicated as the winner (thus it is given a frequency of 1, and all other candidates a frequency of 0).

(12) /bejd, Pret, un/

| /bejd, Pret, un/ Pres: [bejd-] | | | *SPRHVY | OCP-σ | RM | ∃-Integrity | ∃-Ident(+back) | ∃-Ident(-long) | ∃-Max-V | ∃-Ident(-back) | *Pret: $\sqrt{[e]}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| a. | ☞ | bebidun | | | | 1 | | | 1 | 1 | |
| b. | >> | bidun | | | | | | | 1 | 1 | |
| c. | | bebejdun | | | | 1 | | | | | 1 |
| d. | | bejdun | | | 1 | | | | | | 1 |
| e. | | be:jdun | 1 | | | 1 | | 1 | | | |

· In the absence of any learned rankings/weightings, candidates b. and c. harmonically bound candidate a., the "adult" form.

5\. 2. 3. Constrast a tableaux of violations for /geb, Pret, un/:

(13) /geb, Pret, un/

| /geb, Pret, un/ Pres: [geb-] | | | *SPRHVY | OCP-σ | RM | ∃Integrity | ∃-Ident(+back) | ∃-Ident(-long) | ∃Max-V | ∃Ident(-back) | *Pret: $\sqrt{[e]}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| a. | ☞ >> | ge:bun | | | | | | 1 | | | |
| b. | | gegbun | | 1 | | 1 | | | | | |
| c. | | gebun | | | 1 | | | | | | 1 |
| d. | | gegebun | | | | 1 | | | | | 1 |

· Here, candidates b. and d., which violate ∃-Linearity are losers rather than winners, as opposed to the preceding table; b. is also harmonically bounded by d.

5\. 3. Since there is no strict ranking of constraints here that can categorically generate both [ˈbebidun] and *[ˈge:bun], we must use a model of grammar that permits non-categorical outcomes (i.e., variation) and constraint ganging. We will employ a MAXIMUM ENTROPY (MaxEnt) grammar in this function (Goldwater and Johnson 2003, Manning and Schütze 1999: ch. 16).

5. 3. 1. MaxEnt grammars are log linear models that assign probability distributions to a candidate set.

5. 3. 2. A MaxEnt grammar must also be equipped with a learning algorithm that can update constraint weights from some initial search point.

5. 3. 2. 1. The MaxEnt learning models employed in work by Hayes & Wilson (Wilson 2006, Hayes and Wilson 2008, Hayes et al 2009a) employ an off-line (batch) update rule (Conjugate Gradient; Press et al 1992: ch. 10.6)

5. 3. 2. 2. Pater (2008 *et seq.*) has explored the learning of MaxEnt models using on-line update rules (which may now have a convergence proof Boersma and Pater 2013. In contrast to batch learning, here learners update constraint weights learning datum by learning datum.

(14) Perceptron Update Rule (after Pater 2008: 339)
Add $n(x - y)$ to the value of every constraint where $0 < n < 1$, $x =$ the loser's violation marks, and $y =$ the winner's violation marks

5. 4. Off-line Learning Tests (Maxent Grammar Tool, Hayes et al 2009a)

5. 4. 1. Interesting results obtain depending upon the Gaussian prior that is set (thus encoding bias in the learning of certain constraints or constraint types)

5. 4. 2. With no prior whatsoever ($\mu = 0$, $\sigma^2 = 10000$ for all constraints), the grammar simply can't choose between reduplicated ['bebidun] and non-reduplicated ['bidun]: a weight of 0 is assigned to INTEGRITY (the constraint distinguishing those two candidates), and so each receives .5 probability.

5. 4. 3. With non-zero initial weights and a strong prior for all non-morphological constraints ($\mu = 10$, $\sigma^2 = 0.6$), but a weaker prior for morphological constraints ($\mu = 10$, $\sigma^2 = 10$, for RM and *PRET: $\sqrt{[e]}$), interesting results obtain.

(15) MaxEnt Grammar Tool Weights Learned:

| Constraint | Weight |
|---|---|
| *SPRHVY | 10.2182345 |
| OCP-$\sigma$ | 9.999999872 |
| RM | 12.29545141 |
| *PRET: $\sqrt{[e]}$ | 14.61306015 |
| ∃-INTEGRITY | 7.861434916 |
| ∃-IDENT(+back) | 9.400032528 |
| ∃-IDENT(-long) | 10.17944399 |
| ∃-MAX-V | 9.754168254 |
| ∃-IDENT(-back) | 9.546094552 |

5. 4. 4. With a weaker prior for morphological constraints, their weights climb more quickly. The stronger prior for other constraints prevents the weight of INTEGRITY and MAX-V from collapsing too quickly.

5. 4. 5. Under these conditions, the candidates that ultimately correspond to Gothic forms are the most probable, e.g.,

(16) /bejd, Pret, un/

| /bejd, Pret, un/ Pres: [bejd-] | | | *SprHvy | OCP-σ | RM | ∃-Integrity | ∃-Ident(+back) | ∃-Ident(-long) | ∃-Max-V | ∃-Ident(-back) | *Pret. $\sqrt{[e]}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| a. | ☞ 0 | bebidun | | | | 1 | | | 1 | 1 | |
| b. | >> .72 | bidun | | | | | | | 1 | 1 | |
| c. | .03 | bebejdun | | | | 1 | | | | | 1 |
| d. | 0 | bejdun | | | 1 | | | | | | 1 |
| e. | .24 | be:jdun | 1 | | | 1 | | 1 | | | |

5\. 4. 6. Likewise, singular [ˈbajd] takes .99 of the probability distribution, while original "winner" [ˈbe-bajd] receives an insignificant portion.

5\. 4. 7. Obtaining an effect based on bias is similar to bias against "unnatural" phonological constraints reported in Hayes et al 2009b or against "saltatory" mappings in White 2013

5\. 5. On-line Learning Tests (MaxEnt Perceptron)

5\. 5. 1. Boersma and Levelt (2000) showed that the process of constraint weighting using the Gradual Learning Algorithm matched the acquisition more marked syllable structures in Dutch.

5\. 5. 1. 1. Initially high-ranking markedness constraints cause repairs to more marked structures at first, but eventually exposure to the data allows also the marked structures to be produced faithfully.

5\. 5. 1. 2. The order in which the more marked structures are acquired in simulation match the attested acquisition order for Dutch children reported by Levelt and van de Vijver (2004).

5\. 5. 2. Using Praat (v. 5.4), one can run efficient MaxEnt Perceptron simulations:

- "Set decision strategy" as `ExponentialMaximumEntropy`
- When learning, the `Symmetric all` update rule is the update rule given in (22) above.

5\. 5. 3. Using the same constraints and violations used for the off-line tests (all constraints set with an initial disharmony of 100), we used the following settings, suggested by Pater (2009):

- `Evaulation noise`: 0
- `Initial plasticity`: .01
- `Plasticity decrement`: 1
- `Number of plasticities`: 1
- `Rel. placticity spreading`: 0
- `Number of chews`: 1

5\. 5. 4. At ∼ less than 6000 replications of learning (so the learner sees each datum about 6000 times, updating weights with each occurrence), the winner, when evaluated without noise, is the diachronically expected result everywhere except Class VI ([ˈdedabun] continues to win over desired [da:bun]): [ˈbidun] wins over [ˈbebidun], [ˈbajde] wins over [ˈbebajde], etc. The output distributions give a categorical result.

5. 5. 5. Beyond 6000 replications, the older "winners" like [ˈbebidun] begin to receive some winners; eventually, the weights converge on a solution just like the off-line model with no prior.

5. 5. 5. 1. This result looks like another instance of an acquisition order effect: because [ˈbidun] and [ˈbajde] harmonically bound reduplicated [ˈbebidun] and [ˈbebajde], it takes a fair amount of learning for the weight of ∃-Linearity to fall low enough to be irrelevant.

5. 5. 5. 2. As long as all faithfulness constraints are assumed to begin learning with the same non-zero weight, this effect will obtain for these data.

5. 6. Either an interpretation as **acquisition order** or **learning bias** can produce the historically expected results. But the $\sigma^2$ settings here are largely *ad hoc*, so other external evidence for bias towards faster learning/higher weighting of morphological constraints needs to be established.

5. 7. A difficult question: at what age do learners decide to ignore their elders? Do they ever? If maximizing the probability of the observed data will take a very long time or produce unsatisfactory results, do learners abandon hope at some point? Or do they decide that their peers know better?

5. 8. Further work: modeling acquisition of URs, to understand better why [ˈge:bun] isn't parsed as /Red, geb, un/, as mentioned above.

## References


Albright, Adam. 2002. The Identification of Bases in Morphological Paradigms. Ph.D. diss., University of California, Los Angeles.

Albright, Adam, and Bruce Hayes. 2003. Rules vs. Analogy in English Past Tenses: A Computational/Experimental Study. *Cognition* 90.119–161.

Bennett, William H. 1965 [1980]. *An Introducton to the Gothic Language*. New York: Modern Language Association of America.

Boersma, Paul, and Clara Levelt. 2000. Gradual Constraint-Ranking Learning Algorithm Predicts Acquisition Order. In *Proceedings of Child Language Research Forum 30*, 229–37. Stanford: CSLI Publications.

Boersma, Paul, and Joe Pater. 2013. Convergence Properties of a Gradual Learning Algorithm for Harmonic Grammar. `http://www.fon.hum.uva.nl/paul/papers/SGAproof71.pdf`.

Booij, Geert E. 2010. *Construction Morphology*. Oxford: Oxford University Press.

Crosswhite, Katherine. 1999. Intra-paradigmatic Homophony Avoidance in Two Dialects of Slavic. In Matthew Gordon (ed.), *UCLA Working Papers in Linguistics 1: Papers in Phonology*, 48–67. Los Angeles: UCLA Department of Linguistics.

———. 2001. *Vowel Reduction in Optimality Theory*. New York and London: Routledge.

Goldwater, Sharon, and Mark Johnson. 2003. Learning OT Constraint Rankings Using a Maximum Entropy Model. In Jennifer Spenader, Anders Eriksson and Osten Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–20.

Hayes, Bruce, and Colin Wilson. 2008. A Maximum Entropy Model of Phonotactics and Phonotactic Learning. *Linguistic Inquiry* 39.379–440.

Hayes, Bruce, Colin Wilson, and Ben George. 2009a. Manual for Maxent Grammar Tool. `http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool/`.

Hayes, Bruce, Kie Zuraw, Péter Siptár, and Zsuzsa Cziráky Londe. 2009b. Natural and Unnatural Constraints in Hungarian Vowel Harmony. *Language* 85.822–63.

Ito, Junko, and Armin Mester. 2004. Morphological Constrast and Merger: *ranuki* in Japanese. *Journal of Japanese Linguistics* 20.1–18.

Kenstowicz, Michael. 2002. Paradigmatic Uniformity and Constrast. In *Phonological Answers* (*and their Corresponding Questions*), *MIT Working Papers in Linguistics*, vol. 42, 141–63. MIT Department of Linguistics.

Kurisu, Kazutaka. 2001. The Phonology of Morpheme Realization. Ph.D. diss., University of California, Santa Cruz.

Lambdin, Thomas O. 2006. *An Introduction to the Gothic Language*. Eugene, OR: Wipf & Stock.

Levelt, Clara, and Ruben van de Vijver. 2004. Syllable Types in Cross-linguistics and Developmental Grammars. In Rene Kager, Joe Pater and Wim Zonneveld (eds.), *Constraints in Phonological Acquisition*, 204–18. Cambridge: Cambridge University Press.

Manning, Christopher D., and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.

Meid, Wolfgang. 1971. *Das germanische Praeteritum. Indogermanische Grundlangen und Ausbreitung im Germanischen*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Pater, Joe. 2008. Gradual Learning and Convergence. *Linguistic Inquiry* 39.334–45.

———. 2009. Learning in Praat: The Basics. http://people.umass.edu/pater/praat-learning-guide-jp.pdf.

Press, William, Saul Teukolsky, William Vetterling, and Brian Flannery. 1992. *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn. Cambridge: Cambridge University Press.

Struijke, Caro. 2002. *Existential Faithfulness. A Study of Reduplicative TETU, Feature Movement, and Dissimilation*. New York and London: Routledge.

White, James Clifford. 2013. Bias in Phonological Learning: Evidence from Saltation. Ph.D. diss., University of California, Los Angeles.

Wilson, Colin. 2006. Learning Phonology with Substantive Bias: An Experimental and Computational Study of Velar Palatalization. *Cognitive Science* 30.945–82.