# Phonological complexity and input optimization

Michael Hammond
U. of Arizona

## A. Overview

(1) Two points:

a. Input representations are skewed so as to *minimize* constraint violations in the phonology.
b. Input representations are skewed so as to *increase* morphological contrasts.

(2) Organization:

a. overview
b. Rhythm Rule
c. Nasal Assimilation
d. phonological proposal
e. phonological analyses
f. Welsh: Mutation
g. Welsh: Devoicing
h. morphological proposal
i. conclusion

## B. Structure of the proposal

(3) Lexical frequency effects (Fidelholtz, 1975; Hooper, 1976):

| | More frequent | Less frequent |
|---|---|---|
| Reduction: | *astronomy* [ə̲stránəmi] | *gastronomy* [gæ̲stránəmi] |
| Syncope: | *memory* [mɛ́mr̲i] | *mammory* [mǽmə̲ri] |

(4) Phonotactic frequency effects (Davis, 1989; Berkley, 1994):

| | CVC | sCVC |
|---|---|---|
| labial | pope [pop] | — |
| coronal | tight [tajt] | state [stet] |
| dorsal | cake [kek] | — |

(5) Hypothetical: coronal nasal consonants assimilate in place of articulation to a following consonant:

| on pi | an ba | un bo | en do | on ta | un ti | an ku | in ga | on ke |
|---|---|---|---|---|---|---|---|---|
| ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ |
| om pi | am ba | um bo | en do | on ta | un ti | aŋ ku | iŋ ga | oŋ ke |

(6) Input optimization:
Choose input forms that minimize the total number of constraint violations in winning candidates.

(7) More optimized:

| on pi | an ba | en do | on ta | un ti | an ku | in ga |
|---|---|---|---|---|---|---|
| ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ |
| om pi | am ba | en do | on ta | un ti | aŋ ku | iŋ ga |

## C. English: the Rhythm Rule

(8) Rhythm Rule (Liberman & Prince, 1977; Hayes, 1984; Hammond, 1988, etc.)
Stress is shifted to the left in English when the distribution of stresses is sub-optimal.

| | |
|---|---|
| thìrtéen | thírtèen mén |
| Tènnessée | Ténnessèe Tím |
| abrúpt | *ábrùpt péople |

(9) "Too close" does not have to be immediately adjacent:

| | |
|---|---|
| a sléeping adúlt | *an asléep adúlt |
| a blázing inférno | *an abláze inférno |

(10) Bolinger (1962) observes that modifiers with similar meanings are distributed on the basis of their stress properties, essentially to minimize cases where stresses are too close.

| | | |
|---|---|---|
| a sléeping fígure | he was asléep | *an asléep fígure |
| a blázing fíre | it was abláze | *an abláze fíre |

(11) Confirmation:

| | | |
|---|---|---|
| she is hàlf-asléep | hálf-aslèep chíld | *áslèep chíld |
| he is hàlf-alíve | a hálf-alìve chíld | *álìve chíld |

(12) Conjoined adjectives:

| | |
|---|---|
| frèe and éasy | *éasy and frée |
| brìght and shíning | *shíning and bríght |
| gày and yóuthful | *yóuthful and gáy |
| dìre and dréadful | *dréadful and díre |

(13) Hence: adjectives distribute themselves to reduce disfavored stress configurations.

(14) This seems true anecdotally, but is it true statistically?

a. Use the Brown corpus (Kučera & Francis, 1967)
b. Use the CMU pronouncing dictionary (Weide, 1998)
c. Find all disyllabic adjectives in Brown
d. Calculate how frequent different stress patterns are in attribute vs. predicative position

(15) Results:

a. There are 64028 adjective tokens in the 1161192 tagged items in Brown.
b. The CMU dictionary contains 127008 words.
c. The 64028 adjective tokens represent 7521 adjective types.
d. Of these, 4239 occur in the CMU dictionary.
e. Of these, 1340 are disyllabic.

(16)

| Pattern | Example | Types | Tokens | Token freq. |
|---|---|---|---|---|
| σ́σ̆ | happy [hǽpi] | 1002 | 18900 | 0.87 |
| σ̆σ́ | aloof [əlúf] | 177 | 1998 | 0.09 |
| σ́σ̀ | retail [rítèl] | 93 | 678 | 0.03 |
| σ̀σ́ | unfit [ùnfít] | 27 | 262 | 0.01 |

(17)

| Pattern | Example | Non-prenom. tokens | Prenom. tokens |
|---|---|---|---|
| $\acute{\sigma}\breve{\sigma}$ | happy [hǽpi] | 7081 | 11819 |
| $\breve{\sigma}\acute{\sigma}$ | aloof [əlúf] | 997 | 1001 |
| $\acute{\sigma}\grave{\sigma}$ | retail [rítèl] | 171 | 507 |
| $\grave{\sigma}\acute{\sigma}$ | unfit [ùnfít] | 118 | 144 |

(18) The overall distributions in prenominal vs. non-prenominal position are significantly different: $\chi^2(3, N = 8367) = 310.588$, $p < .001$.

(19) No secondary stress: *háppy* and *alóof*. In prenominal position, words like *aloof* represent 8% of adjectives with no secondary stress, while in non-prenominal position, they account for 12%: $\chi^2(1, N = 8078) = 230.695$, $p < .001$.

(20) With secondary stress: *rétàil* vs. *ùnfít*. In prenominal position, words like *unfit* represent 22% of adjectives with secondary stress, while in non-prenominal position, they account for 41%: $\chi^2(1, N = 289) = 58.731$, $p < .001$.

(21) Conjoined adjectives:

a. There are 1520 adjective conjunctions in the Brown corpus.

b. 1326 pairs occur in the CMU pronouncing dictionary.

c. Code each adjective for how far the primary stress is from the right edge.

(22) Coding of conjoined adjectives:

| | $\acute{\sigma}$ | $\acute{\sigma}\sigma$ | $\sigma\acute{\sigma}$ | ... |
|---|---|---|---|---|
| $\acute{\sigma}$ | bíg and táll (1-1) | bíg and hárdy (1-2) | bíg and inért (1-1) | |
| $\acute{\sigma}\sigma$ | háppy and táll (2-1) | háppy and hárdy (2-2) | háppy and inért (2-1) | |
| $\sigma\acute{\sigma}$ | abrúpt and táll (1-1) | abrúpt and hárdy (1-2) | abrúpt and inért (1-1) | |
| ... | | | | |

(23) Means and standard deviations:

| | Mean | St. Dev. |
|---|---|---|
| First Adjective | 1.931 | 0.832 |
| Second Adjective | 2.179 | 0.821 |

(24) This difference is significant: $t(1325) = -9.418, p < .001$, paired.

(25) Conclusions thus far:

a. Bolinger was right;

b. the effect is not (exclusively) lexical;

c. clash is avoided, not the Rhythm Rule per se;

d. we see the effect with singleton adjectives and with conjoined adjectives.

## D. English: Nasal Assimilation

(26)

| | | |
|---|---|---|
| nV | articulate [àrtíkjələt] | inarticulate [ìnàrtíkjələt] |
| | opportune [òpərtún] | inopportune [ìnòpərtún] |
| | efficient [ìfíʃənt] | inefficient [ìnìfíʃənt] |
| mp | pure [pjúr] | impure [ìmpjúr] |
| | probable [prábəbəl] | improbable [ìmprábəbəl] |
| | politic [pálətɪk] | impolitic [ìmpálətɪk] |
| nt | tolerable [tálərəbəl] | intolerable [ìntálərəbəl] |
| | tractable [trǽktəbəl] | intractable [ìntrǽktəbəl] |
| | tangible [tǽnʤəbəl] | intangible [ìntǽnʤəbəl] |
| ŋk | curable [kjúrəbəl] | incurable [ìŋkjúrəbəl] |
| | correct [kərɛ́kt] | incorrect [ìŋkərɛ́kt] |
| | complete [kəmplít] | incomplete [ìŋkəmplít] |

(27) Frequency and stress effects with [k]:

| | Low | High |
|---|---|---|
| Stressed | in-cálculable | in-cápable |
| | in-cáutious | in-còhérent |
| Stressless | in-cŏnsíderate | in-cŏmpléte |
| | in-cŏnspícuous | in- cŏrréct |

(28) Paucity of voiced non-coronals:

| | | |
|---|---|---|
| mb | ? | |
| nd | definite [dɛ́fənət] | indefinite [ìndɛ́fənət] |
| | discreet [dìskrít] | indiscreet [ìndìskrít] |
| | dependent [dəpɛ́ndənt] | independent [ìndəpɛ́ndənt] |
| ŋg | glorious [glóriəs] | inglorious [ìŋglóriəs] |

(29) Liquids (synchronic?):

| | | |
|---|---|---|
| r | relevant [rɛ́ləvənt] | irrelevant [ìrɛ́ləvənt] |
| | reversible [rəvŕ̩səbəl] | irreversible [ìrəvŕ̩səbəl] |
| | rational [rǽʃənəl] | irrational [ìrǽʃənəl] |
| l | legal [lígəl] | illegal [ìlígəl] |
| | licit [lísɪt] | illicit [ìlísɪt] |
| | logical [láʤəkəl] | illogical [ìláʤəkəl] |

(30) Input /n/:

/n+V/ $\Rightarrow$ [nV]
/n+p/ $\Rightarrow$ [mp]
/n+t/ $\Rightarrow$ [nt]
/n+k/ $\Rightarrow$ [ŋk]
/n+l/ $\Rightarrow$ [l]
/n+r/ $\Rightarrow$ [r]

(31) Relative frequency of voiceless labials and dorsals in the Brown corpus:

| | stem | stem freq. | in+stem | in+stem freq. |
|---|---|---|---|---|
| labial | 4732 | 0.48 | 622 | 0.44 |
| coronal | 1443 | 0.15 | 673 | 0.48 |
| dorsal | 3613 | 0.37 | 110 | 0.08 |

(32) The overall difference in the distributions between unprefixed and prefixed is significant: $\chi^2(2, N = 1405) = 1374.571, p < .001$.

(33) Voiceless coronals vs. non-coronals:

| | stem | stem freq. | in+stem | in+stem freq. |
|---|---|---|---|---|
| coronal | 1443 | 0.15 | 673 | 0.48 |
| non-coronal | 8345 | 0.85 | 732 | 0.52 |

(34) This difference is significant as well: $\chi^2(1, N = 1405) = 1228.976$, $p < .001$.

(35) Voiced stops only:

| | stem | stem freq. | in+stem | in+stem freq. |
|---|---|---|---|---|
| voiced coronal | 1021 | 0.22 | 130 | 0.96 |
| voiced non-coronal | 3709 | 0.78 | 5 | 0.04 |

(36) This difference is significant as well: $\chi^2(2, N = 135) = 445.368$, $p < .001$

(37) Why do we get an effect with voiced stops?

a. This difference is simply accidental.

b. There is assimilation with voiced stops with input skewing. That input skewing is great enough to obscure *direct* evidence for the alternation!

(38) Liquids vs. non-liquids:

| | stem | stem freq. | in+stem | in+stem freq. |
|---|---|---|---|---|
| liquid | 5982 | 0.093 | 194 | 0.064 |
| non-liquid | 58046 | 0.907 | 2835 | 0.936 |

(39) This difference is significant as well: $\chi^2(1, N = 3029) = 30.870$, $p < .001$

(40) Are hetero-organic sequences avoided or is nasal assimilation avoided?

| | | |
|---|---|---|
| ram [ræm] | rams [ræmz] | *[rænz] |
| dune [dun] | dunes [dunz] | |
| song [sɔŋ] | songs [sɔŋz] | *[sɔnz] |
| ram [ræm] | rammed [ræmd] | *[rænd] |
| dun [dʌn] | dunned [dʌnd] | |
| bang [bæŋ] | banged [bæŋd] | *[bænd] |

(41) Stem-final nasals in nouns and the plural:

| | stem | stem freq. | stem+s | stem+s freq. |
|---|---|---|---|---|
| labial | 7964 | 0.23 | 2527 | 0.235 |
| coronal | 22585 | 0.652 | 6869 | 0.64 |
| dorsal | 4104 | 0.118 | 1336 | 0.124 |

(42) Coronal vs. non-coronal:

| | stem | stem freq. | stem+s | stem+s freq. |
|---|---|---|---|---|
| coronal | 22585 | 0.652 | 6869 | 0.64 |
| non-coronal | 12068 | 0.348 | 3863 | 0.359 |

(43) This difference is significant as well, but it goes in the *opposite* direction: $\chi^2(1, N = 10732) = 6.471$, $p = 0.011$. Hence, hetero-organic sequences are *not* avoided generally.

(44) Stem-final nasals in verbs and the past:

| | stem | stem freq. | stem+ed | stem+ed freq. |
|---|---|---|---|---|
| labial | 1364 | 0.348 | 625 | 0.286 |
| coronal | 2258 | 0.576 | 1453 | 0.665 |
| dorsal | 296 | 0.076 | 108 | 0.049 |

(45) Coronal vs. non-coronal with verbs:

| | stem | stem freq. | stem+ed | stem+ed freq. |
|---|---|---|---|---|
| coronal | 2258 | 0.576 | 1453 | 0.665 |
| non-coronal | 1660 | 0.424 | 733 | 0.335 |

(46) This difference is significant as well, but it goes in the *usual* direction: $\chi^2(1, N = 2186) = 69.913$, $p < .001$.

(47) Conclusions:

a. there is input skewing with nasal assimilation and deletion as well;
b. coronals are not over-represented in hetero-organic sequences with fricatives;
c. coronals are over-represented in hetero-organic sequences with stops;
d. /n/ does not vacuously assimilate to coronals;
e. liquid assimilation/deletion is productive.

## E. The proposal

(48) **Lexicon Optimization** (Prince & Smolensky, 1993; McCarthy & Prince, 1993)
Suppose that several different inputs $I_1, I_2, \ldots, I_n$ when parsed by a grammar $G$ lead to corresponding outputs $O_1, O_2, \ldots, O_n$, all of which are realized as the same phonetic form Φ—these inputs are all *phonetically equivalent* with respect to $G$. Now one of these outputs must be the most harmonic, by virtue of incurring the least significant violation marks: suppose this optimal one is labelled $O_k$. Then the learner should choose, as the underlying form for Φ, the input $I_k$.

(49) Hypothetical English nasal place assimilation required at the expense of input–output faithfulness.

| | /ɪn-pjur/ | NAS | IO-FAITH |
|---|---|---|---|
| ☞ | ɪmpjur | | ! |
| | ɪnpjur | *! | |

(50) Several possible inputs—/lʌmp/ and /lʌnp/—can produce the same output.

| | /lʌmp/ | NAS | IO-FAITH |
|---|---|---|---|
| ☞ | lʌmp | | |
| | lʌnp | *! | |

| | /lʌnp/ | NAS | IO-FAITH |
|---|---|---|---|
| ☞ | lʌmp | | * |
| | lʌnp | *! | |

(51) Prince & Smolensky propose that we choose the "best" input, the one whose winning output has fewer violations.

| | [lʌmp] | NAS | IO-FAITH |
|---|---|---|---|
| ☞ | /lʌmp/ | | |
| | /lʌnp/ | | * |

(52) **Phonological complexity**
The surface/output forms of a language comprise a set $O = \{O_1, O_2, \ldots, O_n\}$. Every member of that set has a corresponding (optimal) input form $I = \{I_1, I_2, \ldots, I_n\}$. There is, for any phonology, a finite sequence or vector of constraints $C = \langle C_1, C_2, \ldots, C_n \rangle$. Any input–output pairing then defines a vector of violation counts, some number of violations for each constraint earned by the winning candidate for that input. These vectors can, in turn, be summed (constraint by constraint) producing a single vector which is a measure of the phonological complexity of a set of words. To produce a relative measure of phonological complexity given some set of $n$ forms, we divide by $n$.

(53) A less optimal input distribution:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| on pi | an ba | un bo | en do | on ta | un ti | an ku | in ga | on ke |
| ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ |
| om pi | am ba | um bo | en do | on ta | un ti | aŋ ku | iŋ ga | oŋ ke |

(54)

| Input | Output | NAS | IO-FAITH |
|---|---|---|---|
| /on pi/ | om pi | | * |
| /an ba/ | am ba | | * |
| /un bo/ | um bo | | * |
| /en do/ | en do | | |
| /on ta/ | on ta | | |
| /un ti/ | un ti | | |
| /an ku/ | aŋ ku | | * |
| /in ga/ | iŋ ga | | * |
| /on ke/ | oŋ ke | | * |
| | | 0 | 6 |

(55) Relative complexity of the system: $\langle 0, 6 \rangle / 9 = \langle 0, 0.67 \rangle$.

(56) A more optimal input distribution:

| | | | | | | |
|---|---|---|---|---|---|---|
| on pi | an ba | en do | on ta | un ti | an ku | in ga |
| ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ | ⇓ |
| om pi | am ba | en do | on ta | un ti | aŋ ku | iŋ ga |

(57)

| Input | Output | NAS | IO-FAITH |
|---|---|---|---|
| /on pi/ | om pi | | * |
| /an ba/ | am ba | | * |
| /en do/ | en do | | |
| /on ta/ | on ta | | |
| /un ti/ | un ti | | |
| /an ku/ | aŋ ku | | * |
| /in ga/ | iŋ ga | | * |
| | | 0 | 4 |

(58) Relative complexity of this system: $\langle 0, 4 \rangle / 7 = \langle 0, 0.57 \rangle$.

(59) **Input Optimization**
All else being equal, phonological inputs are selected that minimize the phonological complexity of the system.

## F. Analysis: Rhythm

(60) **[FEET]**
A block of constraints that put feet in the right places.

(61) *CLASH
Primary stresses cannot be too close in a phrase.

(62) E-R
Primary stress falls on the rightmost foot.

(63) Words must be stressed:

| | happy | **[FEET]** | *CLASH | E-R |
|---|---|---|---|---|
| ☞ | háppy | | | |
| | happy | *! | | |
| | happý | *! | | |

(64) Rightmost foot has primary stress:

| | Minnesota | **[FEET]** | *CLASH | E-R |
|---|---|---|---|---|
| ☞ | Mìnnesóta | | | |
| | Minnesota | *!* | | |
| | Minnesóta | *! | | |
| | Mìnnesota | *! | | |
| | Mínnesòta | | | *! |

(65) Resolvable stress configuration:

| | thirteen men | **[FEET]** | *CLASH | E-R |
|---|---|---|---|---|
| ☞ | thírtèen mén | | | * |
| | thìrtéen mén | | *! | |
| | thirteen men | *!** | | |
| | thìrtéen men | *! | | |
| | thirteen mén | *!* | | |

(66) Irresolvable stress configuration:

| | aloof man | **[FEET]** | *CLASH | E-R |
|---|---|---|---|---|
| ☞ | alóof mán | | * | |
| | áloof mán | *! | | |
| | álòof mán | *!* | | * |
| | alóof man | *!* | | |
| | aloof mán | *! | | |

(67)

| Input | Output | **[FEET]** | *CLASH | E-R | |
|---|---|---|---|---|---|
| happy man | háppy mán | | | | |
| aloof man | alóof mán | | * | | ⟵ minimize |
| retail man | rétàil mán | | | | |
| unfit man | únfìt mán | | | * | ⟵ minimize |

## G. Analysis: Nasal Assimilation

(68) FAITH-MANNER (F(MAN))
The values for the manner features cannot change.

(69) FAITH-OBS (F(OBS))
Obstruents can’t change *any* features.

(70) *HETERO-NASAL-STOP (HNS)
A nasal must be homo-organic with a following stop.

(71) FAITH-POA (F(POA))
Place of articulation cannot be changed.

(72) Resolvable sequences:

| | i/n-p/ossible | F(MAN) | F(OBS) | HNS | F(POA) |
|---|---|---|---|---|---|
| ☞ | i[mp]ossible | | | | * |
| | i[np]ossible | | | *! | |
| | i[nt]ossible | | *! | | |
| | i[dp]ossible | *! | | | |
| | i[nf]ossible | *! | | | |

(73) Prevocalic nasal:

| | i/n/ordinate | F(MAN) | F(OBS) | HNS | F(POA) |
|---|---|---|---|---|---|
| ☞ | i[n]ordinate | | | | |
| | i[m]ordinate | | | | *! |

(74) With a coronal stop:

| | i/n/tolerant | F(MAN) | F(OBS) | HNS | F(POA) |
|---|---|---|---|---|---|
| ☞ | i[n]tolerant | | | | |
| | i[m]tolerant | | | *! | * |

(75) Why is there no assimilation stem-finally, e.g. in words like *seemed* [simd], *[sind]?

(76) FAITH-STEM (F(STEM))
Stem-final changes are dispreferred. (ANCHOR(R)?)

(77) Stem-final nasals:

| | see/md/ | F(MAN) | F(STEM) | F(OBS) | HNS | F(POA) |
|---|---|---|---|---|---|---|
| ☞ | see[md] | | | | * | |
| | see[nd] | | *! | | | * |
| | see[mb] | *! | | | | * |

(78)

| Input | Output | ... | HNS | F(POA) | |
|---|---|---|---|---|---|
| i/n-p/ossible | i[mp]ossible | | | * | ⟵ minimize |
| i/n-t/olerant | i[nt]olerant | | | | |
| i/n-k/oherent | i[ŋk]oherent | | | * | ⟵ minimize |
| i/n/ordinate | i[n]ordinate | | | | |
| see/md/ | see[md] | | * | | ⟵ minimize |
| du/nd/ | du[nd] | | | | |
| ba/ŋd/ | ba[ŋd] | | * | | ⟵ minimize |

## H. Welsh: Mutation

(79)

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Mae | fy | ngwraig | yn | hoff | o | gathod | a | chŵn. |
| is | my | wife | pred. | fond | of | cats | and | dogs |

‘My wife is fond of cats and dogs.’

(80)

| unmutated | | mutated | | |
|---|---|---|---|---|
| gwraig | [gʷrajg] | ngwraig | [ŋʷrajg] | nasal mutation |
| cathod | [kaθɔd] | gathod | [gaθɔd] | soft mutation |
| cŵn | [kuːn] | chŵn | [xuːn] | aspirate mutation |

(81) Basic Welsh consonant mutations:

| Input | Soft | Nasal | Aspirate |
|---|---|---|---|
| p | b | m̥ | f |
| t | d | n̥ | θ |
| k | g | ŋ̊ | χ |
| b | v | m | n/a |
| d | ð | n | n/a |
| g | ∅ | ŋ | n/a |
| m | v | n/a | n/a |
| ɬ | l | n/a | n/a |
| r̥ | r | n/a | n/a |

(82) The changes are phonological, but the environments are morphosyntactic.

(83) Most productive environments for the Nasal mutation:
a. after *fy* ‘my’;
b. after *yn* ‘in’.

(84) Most productive environments for the Aspirate mutation:
a. after *ei* ‘her’;
b. after *a* ‘and’;
c. after *â/gyda/efo* ‘with’;
d. after *tri* ‘three’;
e. after *chwe(ch)* ‘six’.

(85) Most productive environments for the Soft mutation:
a. feminine singulars after *y* ‘the’ or *un* ‘one’;
b. nouns or adjectives after predicate marker *yn*;
c. adjectives after *mor* ‘so’, *rhy* ‘too’, or *pur* ‘really’;
d. adjectives used with feminine singular nouns;
e. after *dau*/*dwy* ‘two’;
f. after certain prepositions, i.e. *am* ‘for’, *ar* ‘on’, *at* ‘to’, *dan* ‘under’, *dros* ‘over’, *drwy* ‘through’, *heb* ‘without’, *hyd* ‘until’, *gan* ‘by’, *wrth* ‘from’, *i* ‘to’, *o* ‘of’;
g. after *dy* ‘your’ or *ei* ‘his’;
h. direct object of an inflected verb;
i. inflected verbs;
j. nouns after most prenominal adjectives.

(86) Minimal pairs establish that the environment isn't phonological:

| | | |
|---|---|---|
| ei botel | [ibɔtɛl] | 'his bottle' |
| ei photel | [ifɔtɛl] | 'her bottle' |
| eu potel | [ipɔtɛl] | 'their bottle' |
| ei dôn | [idoːn] | 'his tone' |
| ei thôn | [iθoːn] | 'her tone' |
| eu tôn | [itoːn] | 'their tone' |
| ei gar | [igar] | 'his car' |
| ei char | [ixar] | 'her car' |
| eu car | [ikar] | 'their car' |

(87) If mutation is a phonological process, we expect it to be avoided.
Data from the CEG (Ellis et al., 2001) corpus:

| | prep. | prep. freq. | non-prep. | non-prep. freq. |
|---|---|---|---|---|
| mutators | 30405 | 0.31 | 239108 | 0.212 |
| non-mutators | 67779 | 0.69 | 886209 | 0.788 |

(88) The difference goes in the *opposite* direction and is significant: $\chi^2(1, N = 98184) = 5542.824$, $p < .001$.

(89) Proper names of individuals don't undergo mutation:

| | | | |
|---|---|---|---|
| i Pedr | 'to Peter' | i ben | pen 'head' |
| gan Tomos | 'by Thomas' | gan dad | tad 'father' |
| am Catrin | 'about Catherine' | am gi | ci 'dog' |
| at Bethan | 'toward Bethan' | at fws | bws 'bus' |
| heb Dafydd | 'without David' | heb ddyn | dyn 'man' |
| wrth Gerallt | 'to Gerald' | wrth ŵr | gŵr 'husband' |
| o Mair | 'from Mary' | o foch | moch 'pigs' |
| dan Llinos | 'under Llinos' | dan lif | llif 'flood' |
| ar Rhys | 'on Rhys' | ar ran | rhan 'part' |

(90) If mutation is a phonological process, we expect fewer mutating consonants in non-name nouns:

| | non-names | non-names freq. | names | names freq. |
|---|---|---|---|---|
| mutators | 117961 | 0.486 | 6066 | 0.218 |
| non-mutators | 124630 | 0.514 | 21775 | 0.782 |

(91) Again, the difference goes in the *opposite* direction and is significant: $\chi^2(1, N = 27841) = 8027.046$, $p < .001$.

(92) What about in the minimal pair environment? We expect fewer mutating consonants after *ei*:

| | ei | ei freq. | non-ei | non-ei freq. |
|---|---|---|---|---|
| mutators | 9297 | 0.461 | 260216 | 0.216 |
| non-mutators | 10891 | 0.539 | 943097 | 0.784 |

(93) Once again, the difference goes in the *opposite* direction and is significant: $\chi^2(1, N = 20188) = 7107.316$, $p < .001$.

## I. Welsh: Adjectives

(94) Is it that Welsh generally skews in the opposite direction or is it that morphological processes like mutation skew in the opposite direction?

(95) Welsh comparatives/superlatives:

| Stem | | | Comparative | | Superlative | |
|---|---|---|---|---|---|---|
| cyflym | 'fast' | [kəvlɨm] | cyflymach | [kəvləmax] | cyflymaf | [kəvləma(v)] |
| llawn | 'full' | [ɬawn] | llawnach | [ɬawnax] | llawnaf | [ɬawna(v)] |
| hapus | 'happy' | [hapɨs] | hapusach | [hapɨsax] | hapusaf | [hapɨsa(v)] |
| tawel | 'quiet' | [tawɛl] | tawelach | [tawɛlax] | tawelaf | [tawɛla(v)] |
| pell | 'far' | [pɛɬ] | pellach | [pɛɬax] | pellaf | [pɛɬa(v)] |
| doeth | 'wise' | [dɔjθ] | doethach | [dɔjθax] | doethaf | [dɔjθa(v)] |
| araf | 'slow' | [ara(v)] | arafach | [aravax] | arafaf | [arava(v)] |

(96) Voiced stops devoice in this context:

| Stem | | | Comparative | | Superlative | |
|---|---|---|---|---|---|---|
| gwlyb | 'wet' | [gʷlɨb] | gwlypach | [gʷləpax] | gwlypaf | [gʷləpa(v)] |
| caled | 'hard' | [kalɛd] | caletach | [kalɛtax] | caletaf | [kalɛta(v)] |
| parod | 'ready' | [parɔd] | parotach | [parɔtax] | parotaf | [parɔta(v)] |
| enwog | 'famous' | [ɛnʷɔg] | enwocach | [enʷɔkax] | enwocaf | [ɛnʷɔca(v)] |
| hyfryd | 'lovely' | [həvrɨd] | hyfrytach | [həvrətax] | hyfrytaf | [həvrəta(v)] |
| teg | 'fair' | [teːg] | tecach | [tekax] | tecaf | [teka(v)] |
| pwysig | 'important' | [pujsɪg] | pwysicach | [pwjsɪkax] | pwysicaf | [pwjsɪka(v)] |

(97) Stem-medial voiceless stops are rare, but they do occur and do not alternate, confirming a devoicing analysis:

| Stem | | | Comparative | | Superlative | |
|---|---|---|---|---|---|---|
| twp | 'stupid' | [tup] | twpach | [tupax] | twpaf | [tupa(v)] |
| gwyllt | 'wild' | [gʷɨɬt] | gwylltach | [gʷɨɬtax] | gwylltaf | [gʷɨɬta(v)] |
| trist | 'sad' | [trist] | tristach | [tristax] | tristaf | [trista(v)] |
| pert | 'pretty' | [pɛrt] | pertach | [pɛrtax] | pertaf | [pɛrta(v)] |

(98) Corpus analysis:

a. Since voiceless stops are quite rare in this context, we set them aside.
b. Compare voiced stops with nasals.
c. Dorsal nasals are orthographically ambiguous (*ng*), so set dorsals aside and focus on [m,n] vs. [b,d].

(99) Comparison of stem-final voiced stops in adjectives vs. comparatives/superlatives:

| | Adj. | Adj. freq. | Comp./Sup. | Comp./Sup. freq. |
|---|---|---|---|---|
| Voiced stops | 8940 | 0.473 | 72 | 0.351 |
| Nasals | 9946 | 0.527 | 133 | 0.649 |

(100) The difference goes in the expected direction and is significant: $\chi^2(1, N = 205) = 12.269$, $p < .001$. Hence, it is mutation that is reversed, not Welsh generally.

## J. Morphological proposal

(101) An apparent contradiction:

a. for phonological generalizations, constraint violations are *avoided* by skewing input representations;
b. for morphological generalizations, constraint violations are *preferred* by skewing input representations.

(102) [MUTATE] (MUT)
Block of constraints that ensure that initial consonants take on appropriate forms in mutation contexts.

(103) FAITH-IO (F-IO)
Features don't change.

(104) Mutator in mutation context:

| | am /k/i | MUT | F-IO |
|---|---|---|---|
| ☞ | am [g]i | | * |
| | am [k]i | *! | |

(105) Mutator in non-mutation context:

| | /k/i | MUT | F-IO |
|---|---|---|---|
| ☞ | [k]i | | |
| | [g]i | | *! |

(106) Non-mutator regardless of context:

| | (am) afal 'apple' | MUT | F-IO |
|---|---|---|---|
| ☞ | (am) afal | | |

(107)

| Input | Output | MUT | F-IO | |
|---|---|---|---|---|
| am /k/i | am [g]i | | * | ⟵ minimize? |
| am afal | am afal | | | |
| /k/i | [k]i | | | |
| afal | afal | | | |

(108) REALIZE MORPHEME (RM)
Let $\alpha$ be a morphological form, $\beta$ be a morphosyntactic category, and $\mathrm{F}(\alpha)$ be the phonological form from which $\mathrm{F}(\alpha+\beta)$ is derived to express a morphosyntactic category $\beta$. Then RM is satisfied with respect to $\beta$ iff $\mathrm{F}(\alpha+\beta) \neq \mathrm{F}(\alpha)$ phonologically (Kurisu, 2001).

(109)

| Input | Output | RM | F-IO | |
|---|---|---|---|---|
| am /k/i | am [g]i | | * | |
| am afal | am afal | * | | ⟵ minimize |
| /k/i | [k]i | | | |
| afal | afal | | | |

(110) Morphological conclusions:
a. morphological exponence should be uniform;
b. we can get that effect with RM;
c. Input optimization requires constraint ranking and/or weighting.

## K. Conclusion

(111) Input optimization:
phonological complexity measured in terms of constraint violations is minimized.

(112) This mechanism extends naturally to describe frequency effects with phonotactic markedness as well.

(113) Why do less optimal forms *ever* occur? There must be sufficient contrasts available to the system to achieve expressiveness, but marked configurations are restricted as much as possible in use.

(114) Auxiliary results:
   a. the grammar must be constraint-based (though not necessarily OT);
   b. there is rhythm-based skewing with individual adjectives and with conjoined adjectives;
   c. clash and stress shift are both avoided;
   d. there is place-based skewing with *in*- and with stem-final nasals;
   e. there is deletion/degemination-based skewing with *in*-;
   f. there is no skewing with *-s*; hence only pre-stop nasals must assimilate;
   g. there is no vacuous rhythm or assimilation;
   h. there is also input skewing so that morphological exponence is uniform;
   i. there is skewing of voicing in stem-final consonants in Welsh.

(115) Implications:
   a. we have another tool with which we can investigate phonological generalizations;
   b. we can hypothesize how phonology is learned with this mechanism;
   c. we can distinguish morphology from phonology with this tool.

## L. References


Albright, Adam (2009). Feature-based generalization as a source of gradient acceptability. *Phonology* **26**:9–41.

Berkley, Deborah Milam (1994). The OCP and gradient data. *Studies in the Linguistic Sciences* **24**:59–72.

Bolinger, D. (1962). Binomials and pitch accent. *Lingua* **11**:34–44.

Coleman, John & Janet Pierrehumbert (1997). Stochastic phonological grammars and acceptability. In *Computational Phonology: Third meeting of the ACL Special Interest Group in Computational Phonology*. Somerset: Association for Computational Linguistics, 49–56.

Davis, Stuart (1989). Cross-vowel phonotactic constraints. *Computational Linguistics* **15**:109–111.

Ellis, N. C., C. O'Dochartaigh, W. Hicks, M. Morgan & N. Laporte (2001). Cronfa electroneg o Gymraeg (CEG): A 1 million word lexical database and frequency count for Welsh. `http://www.bangor.ac.uk/canolfanbedwyr/ceg.php.en`.

Fidelholtz, James (1975). Word frequency and vowel reduction in English. In *Chicago Linguistic Society*, vol. 11. 200–213.

Frisch, S., N. R. Large & D. B. Pisoni (2000). Perception of wordlikeness: effects of segment probability and length on the processing of nonwords. *Journal of Memory and Langauge* **42**:481–496.

Golston, Chris (1998). Constraint-based metrics. *Natural Language and Linguistic Theory* **16**:719–770.

Graff, Peter (2012). *Communicative Efficiency in the Lexicon*. Ph.D. thesis, MIT.

Hammond, Michael (1988). *Constraining Metrical Theory: A Modular Theory of Rhythm and Destressing*. New York: Garland. 1984 UCLA doctoral dissertation.

Hammond, Michael (1999). *The Phonology of English*. Oxford: Oxford University Press.

Hammond, Michael (2004a). Frequency, cyclicity, and optimality. *Studies in Phonetics, Phonology, and Morphology* **10**:349–364.

Hammond, Michael (2004b). Gradience, phonotactics, and the lexicon in English phonology. *International Journal of English Studies* **4**:1–24.

Hammond, Michael (2011). Welsh mutations and statistical phonotactics. In Andrew Carnie (ed.), *Formal Approaches to Celtic Linguistics*. Newcastle upon Tyne: Cambridge Scholars Publishing, 337–358.

Hammond, Michael (2013). Input optimization in English. *Journal of the Phonetic Society of Japan* **17**:1–12.

Hayes, Bruce (1984). The phonology of rhythm in English. *Linguistic Inquiry* **15**:33–74.

Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* **39**:379–440.

Hayes, Rachel (2003). *How are Second Language Phoneme Contrasts Learned?* Ph.D. thesis, University of Arizona.

Hooper, Joan (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (ed.), *Current progress in historical linguistics*. Amsterdam: North Holland, 96–105.

Kučera, H. & W.N. Francis (1967). *Computational Analysis of Present-day American English*. Brown University Press.

Kurisu, Kazutaka (2001). *The Phonology of Morpheme Realization*. Ph.D. thesis, University of California Santa Cruz.

Liberman, Mark & Alan Prince (1977). On stress and linguistic rhythm. *Linguistic Inquiry* **8**:249–336.

Maye, Jessica (2000). *The Acquisition of Speech Sound Categories on the Basis of Distributional Information*. Ph.D. thesis, University of Arizona.

McCarthy, John & Alan Prince (1993). Prosodic morphology. U. Mass.

Pierrehumbert, Janet B., Jennifer Hay & Mary E. Beckman (2004). Speech perception, well-formedness, and the statistics of the lexicon. In John Local, Richard Ogden & Rosalind Temple (eds.), *Papers in Laboratory Phonology VI*. Cambridge: Cambridge University Press, 58–74.

Prince, Alan & Paul Smolensky (1993). Optimality Theory. U. Mass and U. of Colorado.

Saffran, Jenny R., Richard Aslin & Elissa L. Newport (1996). Statistical learning by 8-month-old infants. *Science* **274**:1926–1928.

Trubetzkoy, Nikolai (1939). *Grundzüge der Phonologie*.

Weide, Robert L. (1998). The CMU pronouncing dictionary. `http://www.speech.cs.cmu.edu/cgi-bin/cmudict`.

Zamuner, Tania (2001). *Input-based Phonological Acquisition*. Ph.D. thesis, U. of Arizona.

Zamuner, Tania, LouAnn Gerken & Michael Hammond (2004). Phonotactic probabilities in young children's speech production. *Journal of Child Language* **31**:515–536.