Morphologically conditioned phonetics? Not proven.
Ricardo Bermúdez-Otero (University of Manchester)

INTRODUCTION

The fundamental philosophical questions like the continuous versus the discrete or the limits of knowledge are never definitively solved.

Gregory Chaitin, Metamats, 2.

§1 In classical modular feedforward architectures of grammar (Pierrehumbert 2002, Bermúdez-Otero 2007), there is no direct communication between morphology and phonetics; morphology can affect the application of gradient processes of phonetic implementation only indirectly, through categorical properties of the surface phonological representation.

§2 So-called English /t,d/-deletion appears to challenge the classical architecture (Myers 1995, Coetzee and Pater 2011):

• morphological sensitivity (e.g. Labov et al. 1968; Guy 1980, 1994)
  /t,d/ are deleted more often when affiliated to a stem mist
  than when affiliated to a regular inflectional suffix miss-ed

• phonetic gradience (Browman and Goldstein 1989, 1990)
  in at least some instances, the tongue-tip gesture of the allegedly deleted /t,d/ is not removed, but merely hidden by adjacent overlapping gestures (and also possibly reduced).

§3 The rate of complete (categorical) vs incomplete (gradient) /t,d/-deletion is currently unknown.

In any case, the classical architecture can describe the facts in §2 by means of a two-step derivation (cf. Scobbie 1995):

(1) phonology: variable, categorical, morphologically sensitive featural or prosodic change
(2) phonetics: variable, gradient, morphologically insensitive gestural adjustment.

§4 Two-step derivations along the lines of §3 should not be regarded as contrived; they are the natural and expected result of the highly prevalent diachronic process of rule scattering.

RULE SCATTERING (after Robinson 1976):
Diachronic change often introduces a new avatar of an existing pattern into a higher component of the grammar by reanalysis, while the old avatar remains in situ.

1 Note added on 17 Dec 2010. Some two weeks after this paper was presented at OnLI II, it was brought to my attention that J. Myers had reached very similar conclusions about English /t,d/-deletion in an entirely independent study, reported in Myers (1995). I have added references to Myers’s manuscript, highlighted in red, at relevant points in this handout.
§5 In classical architectures, there cannot be direct communication between morphology and phonetics because morphology and phonetics do not share an interface; the (categorical) phonology always mediates between the two.

§6 (a) Late 1980s interactionist architecture
(Kiparsky 1985, Booij and Rubach 1987)
(b) Late 1980s noninteractionist architecture
(Halle and Vergnaud 1987, Odden 1993)

§7 A long-standing principle (see Scheer 2010):

- Kruszewski (1881 [1995]: 27):
  'the combinatory process [=coarticulation] cannot be impeded by “analogy”.'

- Structuralism (e.g. Z. S. Harris 1951, ch. 8):
morphosyntax controls the distribution of juncture phonemes; allophony rules refer to juncture phonemes, not to morphosyntax.

- SPE (Chomsky and Halle 1968): a variant of §6(b).

- Bidirectional Phonology and Phonetics (Boersma 2009):
  parallel, rather than feedforward, architecture;
  but, still, no constraints directly evaluating the relationship between Morphemes and Auditory Form or Articulatory Form.
Empirical challenges

§8 Allegations of direct morphological conditioning of gradient phonetic processes (see also Kawahara 2011: §2.3.3):

<table>
<thead>
<tr>
<th>phenomenon</th>
<th>reference</th>
<th>this paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>• English /t,d/-deletion</td>
<td>Myers (1995), Coetzee and Pater (2011)</td>
<td>§22ff</td>
</tr>
<tr>
<td>• English /l/-darkening</td>
<td>Sproat and Fujimura (1993)</td>
<td>†</td>
</tr>
<tr>
<td>• The Withgott Effect</td>
<td>Steriade (2000)</td>
<td>§9-§13</td>
</tr>
<tr>
<td>• Korean gestural timing</td>
<td>Cho (2001)</td>
<td>§18-§21</td>
</tr>
<tr>
<td>• Dutch final devoicing</td>
<td>Ernestus and Baayen (2006, 2007)</td>
<td></td>
</tr>
<tr>
<td>• Dutch -igheid</td>
<td>Pluymaekers et al. (2010)</td>
<td></td>
</tr>
</tbody>
</table>

† I assume that the approach to /t,d/-deletion defended in this paper is also appropriate for /l/-darkening: see Hayes (2000: 93), Bermúdez-Otero (2007: note 6), but specially Bermúdez-Otero (2013: §2.4).

Prosodic solutions (I): the Withgott Effect

§9 The Withgott effect (Withgott 1982):

cápi[r]al ⇔ căpi[r]alístic
míli[tʰ]árly ⇔ míli[tʰ]arístic

The effect is subject to variation (Riehl 2003).

§10 Steriade’s (2000) claims:

• Flapping is a gradient phonetic process (shortening of stop closure); cf. §11 below.
• The Withgott effect requires analogical transfer (via output-output correspondence) of gradient phonetic properties; cf. §12 below.

§11 In fact, flapping is a variable but categorical phrase-level process creating discrete stop vs flap allophones:

• Stop and flap allophones are distinguished by several interacting acoustic cues (Riehl 2003).
• Stop and flap allophones create a bimodal distribution in duration (Herd et al. 2010: 508).
• /t/-flaps and /d/-flaps are phonetically indistinguishable: e.g. same mean duration (Zue and Laferriere 1979: 1043-44, Kwong and Stevens 1999: 3, Herd et al. 2010: 511, Braver 2010).

The underlying /t-/d/ contrast is reflected only in the preceding vowel, notably in its duration. This effect is entirely expected since prefortis clipping (vowel shortening before voiceless obstruents) can be independently shown to apply at the stem level and so to precede phrase-level flapping (Bermúdez-Otero 2004: §14-§22).

§12 In fact, morphological structure affects flapping not directly, but indirectly via prosody:

foot structure is built cyclically
and an intervocalic /t/ resists flapping if it is initial in some foot-projection at the word level.


www.bermudez-otoero.com
§13  Prosodic solutions (II): durational effects of suffixation

§14  Sugahara and Turk (2009) on durational effects of word-level suffixation in Scottish English:

phonetic realization of the string /ɪːz/:
- shorter duration in *raisin* [ɪːːzn̩]
- longer duration in *rais-ing* [ɪːzɪŋ]

The effect is relatively small:
- it was not detected at all in Sproat and Fujimura (1993)
- not statistically significant at ‘normal’ speech rates (Sugahara and Turk 2009: 496)
- 6.6% difference (mean of 23ms) at ‘slow’ speech rates (Sugahara and Turk 2009: 496)
- 9.6% difference (mean of 42ms) at ‘extra-slow’ speech rates (Sugahara and Turk 2009: 499).

§15  A prosodic solution (Bermúdez-Otero 2011: §4):

- stem level (stress-affecting affixation)  ⇒  exhaustive footing
- word level (stress-neutral affixation)  ⇒  stray syllables attached directly to ω

§16  raisin  raise  rais-ing

disyllabic foot  but  monosyllabic foot
Sugahara and Turkel propose the prosodification [\( \omega \rightarrow [\omega \rightarrow \omega] \)] for \textit{rais-ing}, but this incorrectly predicts that the stem will behave durationally like the first member of a compound \citep[\textit{pace} Sugahara and Turkel 2009: 488; cf.][]{Sproat1993}. 

**Processing solutions? A speculative suggestion**

\section{18} Cho (2001) on gestural timing in Korean:
- tautomorphemic segments \( \Rightarrow \) low variability in gestural timing /napi/ ‘butterfly’
- heteromorphemic segments \( \Rightarrow \) high variability in gestural timing /nap-i/ ‘lead-NOM’

I assume that Cho is right in treating Korean nominal inflections as belonging in the same morphological and prosodic word as the stem: cf. Yoon (2005) for a phrasal-affix analysis, but see Bermúdez-Otero and Payne (2011).

\section{19}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure.png}
\caption{(Cho 2001: 131, Figure 1)}
\end{figure}

‼ In this case, the effect of morpheme boundaries is symmetrical: both more extreme overlap and more extreme nonoverlap; cf. §14.

\section{20} Cho’s analysis:

‘[...]he phonological structure of a lexical entry forms a “constellation” of gestures, a stable organization among gestures whose timing is specified in the lexicon’ (p. 153, my emphasis).


\section{21} A speculative suggestion: cascading activation.

Processing models with cascading activation allow competing representations at processing stages to pass on activation to subsequent processing stages even before selection is complete at stage $\S$.
- In such models, competing morphs can pass on activation downwards to phonological representations and thence to articulatory planning even before morph selection is complete.
- If that happens, we expect higher amounts of articulatory ‘jitter’ near points of morph selection (morphological boundaries) than inside morphs.
- In this scenario, articulatory jitter is an artifact of processing: phonological representations can be more or less activated in processing, but phonological representations do not contain fine-grained detail about gestural timing.


### A two-stage derivation of English /d/-deletion

#### Morphological sensitivity

<table>
<thead>
<tr>
<th>§22</th>
<th>Morphological affiliation</th>
<th>Examples</th>
<th>Surface forms</th>
<th>% deleted (Guy 1991a: 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>stem</td>
<td>mist, pact</td>
<td>[ms(t), pæk(t)]</td>
<td>38.1</td>
<td></td>
</tr>
<tr>
<td>regular inflectional suffix</td>
<td>miss-ed, pack-ed</td>
<td></td>
<td>16.0</td>
<td></td>
</tr>
</tbody>
</table>

This correlation between the morphological affiliation of the stop and deletion rate has been replicated in various dialects (e.g. Guy 1994: 140), but at least two studies have warned of possible confounding variables: the preceding consonant (Tagliamonte and Temple 2005: 295-96) and lexical frequency (Guy et al. 2008).

| §23 | Irregular weak verbs (e.g. lef-t, tol-d) display an intermediate average rate of deletion in corpora: e.g. 33.9% in Guy (1991a: 5). This does not reflect a third type of morphological structure (pace Guy 1991a,b), but is rather an effect of age grading (Guy and Boyd 1990, Guy 2010: 255):
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• lef-t, tol-d</td>
<td>behave like mist, pact for younger speakers</td>
</tr>
<tr>
<td>• lef-t, tol-d</td>
<td>behave like miss-ed, pack-ed for older speakers</td>
</tr>
</tbody>
</table>

#### Phonetic gradience

| §24 | Audibly released /t/ in perfect memory (Brownman and Goldstein 1990: 364, figure 19.13(a)) |

![Audio waveform diagram](image-url)
§25 Inaudible /t/ with hidden tongue-blade gesture (τ) in *perfect memory* (Browman and Goldstein 1990: 364, figure 19.13(b))

![Diagram of articulatory gestures](image)

§26

- How many of the instances of /t,d/-deletion reported in the sociolinguistic literature are actually incomplete, like §25?

  We just don’t know.

- Can we assume that /t,d/-deletion is always incomplete (see e.g. Temple 2009)?

  No! Lichtman (2010) examined the Wisconsin Microbeam Database and conducted an additional electromagnetic articulography study, and found that all speakers produced at least some instances of deleted /t/ without residual linguoalveolar gestures.

- Cf. the following cautionary tale:

  Early reports of gradient coarticulation in assimilatory external sandhi led to the belief that categorical feature delinking and spreading did not operate across word-boundaries (e.g. Barry 1985, Wright and Kerswill 1989, Nolan 1992, Hardcastle 1995, Zsiga 1995); but this belief has proved false: categorical assimilatory external sandhi does exist (e.g. Nolan et al. 1996, Ellis and Hardcastle 2002, Kochetov and Pouplier 2008: 414).

**Two-step derivations**

§27 Assume the worst-case scenario: morphologically-sensitive incomplete /t,d/-deletion is prevalent.

Can this be accommodated within the classical architecture?

Yes, by means of two step-derivations:

1. **Phonology:** variable, categorical, morphologically sensitive feature or prosodic change
2. **Phonetics:** variable, gradient, morphologically insensitive gestural adjustment.

See Myers (1995) for a similar view.
§28  *Scobbie’s Caveat* (after Scobbie 1995):

If, as commonly assumed, phonological processes need not be neutralizing (i.e. need not be ‘structure-preserving’ in the sense of Kiparsky 1985), then evidence of gradience is not evidence of absence of a categorical effect.

§29  Key question:

Is there an independently motivated way of categorically representing /t,d/-reduction in the phonology?

§30  • Option 1: represent the absence of audible release

  e.g. in Element Theory, delete the h° element (noise component)

  (e.g. J. Harris and Lindsey 1995)

• Option 2: represent concurrent articulation

  e.g. tuck the underlying /t,d/ under the timing slot of the preceding consonant

  \[ \text{X X X X X X X} \]
  \[ \text{| | | | | | |} \]
  \[ \text{d i } p \text{-d} \rightarrow \text{d i p t} \]

  This autosegmental chart iconically represents

  • preforsis clipping (reduced duration of /i/ before longer fortis /p/)
  • overlap of /p/ and /d/.

**Deriving the effect of morphology on frequency cyclically (Guy 1991a, 1991b)**

§31  One round of variable reduction in *miss-ed*  \( \Rightarrow \) low reduction rate

\[ \text{UR \quad \left[\text{wL, [sl, mis]} \text{ d}\right]} \]

stem level

\[ \text{mis} \]

word level

\[ \text{mis} \]

\[ \text{mst} \]

\[ \text{mst}^* \]

SR

\[ \text{[mst]} \]

\[ \text{[mst}^*] \]
§32 Two rounds of variable reduction in *mist*  \(\Rightarrow\) high reduction rate

UR

\[
\begin{array}{c}
\text{stem level} \\
\text{word level} \\
\text{SR}
\end{array}
\begin{array}{ccc}
\text{mist} & \text{mist}^* \\
\downarrow & \downarrow \\
\text{[mist]} & \text{[mist}*]
\end{array}
\]

§33 *Pace* Guy (1991a, 1991b), I do not assume that the rate of reduction must be the same in every cycle.

On the contrary:
- In their normal diachronic life cycle (Bermúdez-Otero 2007: 503-5, 2013; Bermúdez-Otero and Trousdale 2012: §2), phonological processes enter the grammar from below and climb up to progressively higher levels.
- This predicts the possibility that a process may display higher rates of application at lower levels, where it has been active the longest (e.g. Turton 2012 on /l/-darkening).

TWO-STAGE DERIVATIONS ARE NATURAL: DIACHRONIC RULE-SCATTERING

A doctrine is not judged at all until it is judged in its best form.

*John Stuart Mill, ‘Sedgwick’s Discourse’, quoted in John Rawls, ‘Some remarks about my teaching’*

**Diachronic rule-scattering gives rise to two-step derivations**

§34 A perfectly ordinary stage of affairs:

two cognate rules, one gradient, one categorical, representing different stages in the diachronic life cycle of the same sound pattern, coexist within the same grammar.

Examples:
- Philadelphia \(\varphi\)-tensing (Labov 1994)  §36–§37 below
- Pre-yod /s/-palatalization (Zsiga 1995)  §38–§40 below

§35 Why is this stage of affairs common? Presumably because of the mechanics of *stabilization* (i.e. of the diachronic process of reanalysis whereby tokens of a gradient rule of phonetic implementation are reinterpreted as involving a categorical distinction):

the reanalysed tokens give rise to a new categorical rule, whilst the nonreanalysed tokens preserve the gradient rule *in situ* (see Bermúdez-Otero 2013: example (7) and surrounding discussion).

*Æ* Just a special case of rule scattering: see §4.
Philadelphia æ-tensing

§36  *A categorical default stem-level rule...*

- categorical opposition between lax [æ] and tense [æː]
- the phonetic realizations of [æ] and [æː] occupy largely nonoverlapping regions in auditory space (Labov 1989: 8-10)
- default pattern: æ is normally tense before coda /m, n, f, 0, s/  (Labov 1994: 430)
- counterbled by resyllabification before word-level suffixes:
  e.g. m[æː]n, m[æ]nner, m[æː]nn-ing  (Labov 1994; 430-31)
- outright lexical exceptions: [æː] in mad, bad, glad; but [æ] in sad, fad, lad
  [æ] in strong verbs like began, ran, swam

§37  *...overlaid with a gradient phonetic process*

the precise formant values of the acoustic realization of [æː] are exquisitely sensitive to the phonetic environment

<table>
<thead>
<tr>
<th></th>
<th>effect on height</th>
</tr>
</thead>
<tbody>
<tr>
<td>following nasal</td>
<td>158</td>
</tr>
<tr>
<td>preceding nasal</td>
<td>99</td>
</tr>
<tr>
<td>preceding obstruent plus /l/</td>
<td>−210</td>
</tr>
<tr>
<td>two following syllables</td>
<td>−410</td>
</tr>
<tr>
<td>etc.</td>
<td></td>
</tr>
</tbody>
</table>

(coefficients in Labov 1994: 466)

Pre-yod /s/-palatalization

§38  *A neutralizing stem-level morphophonological process...*

\[\text{confi/s}+\text{/j/}\text{on} \rightarrow \text{confi}[\text{j}]\text{on}\]

[J] realized with same pattern of linguopalatal consonant as underlying /ʃ/

§39  *...coexisting with gradient gestural overlap across word-boundaries*

\[\text{pre/s}\# /\text{j/}\text{ou} \rightarrow \text{pre[ð]}\text{ou}\]

[ð] arises from the superposition of the patterns of linguopalatal contact for /s/ and /j/

§40  But the same diachronic origin! cf. Chaucer’s confi/si.uːn/.
METHODODOLOGICAL CONSIDERATIONS

The insufficiency of conceptual arguments

§41 *Implications of the evidence for rule-scattering* (§34–§40)

Criteria of elegance or simplicity (e.g. Ockham’s Razor) do not suffice to dismiss the two-step solution to the problem that English *t/d*-deletion raises for classical architectures.

Two-step derivations may look synchronically complex, but they are diachronically natural.

Of course, the evidence from rule scattering does not prove that two-step derivations are right for English *t/d*-deletion.

§42 *Conceptual arguments against direct communication between morphology and phonetics*

Grotesquely nonclassical behaviours become conceivable:

- ‘Mark dual number by adding /-no/ and increasing gestural overlap by 20%.’
- ‘Form eventive nominalizations by adding /-ti/ and lengthening VOT by 30%.’

Of course, usage-based models need not predict such aberrant behaviour:

- Bybee’s line (2001: 54):
  Speakers are driven by efficiency requirements to constantly reuse a finite set of highly practised neuromotor programmes in speech. Presumably, this prevents morphological categories from drifting apart in phonetic space.

However, in the absence of equations or computer simulations, isn’t this just a promissory note?

“This family of theories [sc. exemplar-based theories—RBO] holds that the lexicon consists of a massive number of stored phonetic (or auditory) events, with or without category labels. Those subtheories that touch on category creation can do so because they include no category labels, but those subtheories that make interesting linguistic generalizations [...] do require the presence of category labels.” (Boersma 2012: §10.2.7)

Empirical problems

§43 *Limitations of the data currently available*

- sociolinguistic studies: impressionistic coding
  no way of controlling for the presence of hidden gestures
- articulatory studies: few speakers
  few tokens
  lack of control of relevant morphological and phonological factors
§44 Analytical challenges

A bimodal distributions on some phonetic dimension is usually treated as the clearest indication of a categorical distinction (e.g. §11 above):

Problem: suppose that the realizations of two categories create two normal densities on some phonetic dimension;
the mixture of these two densities may be unimodal or very weakly bimodal under a wide range of conditions.

(Schilling et al. 2002: 225, figure 6)

See Myers (1995) for discussion with particular reference to the problem of t/d-deletion.

REFERENCES


---

**CONTACT DETAILS**

Ricardo Bermúdez-Otero
Linguistics and English Language
University of Manchester
Oxford Road
Manchester M13 9PL
United Kingdom
r.bermudez-oter@manchester.ac.uk
www.bermudez-oto.com