Dialectologia 18 (2017), 129-155. ISSN: 2013-2247

Received 5 January 2015. Accepted 23 March 2015.

PHONETIC DETAIL AND VARIATIONIST PHONOLOGY: THE CASE OF (t,d)

Rosalind A. M. TEMPLE

University of Oxford*

rosalind.temple@new.ox.ac.uk

Abstract

"-t,d deletion", or "(t,d)" has been the object of variationist studies for over half a century and continues to play a key part in theoretical debates about phonological representations both within and beyond variationist linguistics. Whereas they differ in their interpretations of the findings, most studies of the variable share a set of key assumptions about the nature of the variation involved. This paper addresses some of those assumptions, demonstrating how an examination of the detailed phonetics of the data raises fundamental problems which suggest that, in the absence of independent evidence to the contrary, (t,d) is best modelled as a Connected Speech Process, albeit a cognitively determined one.

Keywords

-t,d deletion, variation, phonetics, phonology, CSPs

DETALLE FONÉTICO Y FONOLOGÍA VARIACIONISTA: EL CASO DE (t,d)

Resumen

"La caída de -t, d", o "(t,d)" ha sido objeto de estudios variacionistas durante más de medio siglo y continúa desempeñando un papel clave en los debates teóricos sobre las representaciones fonológicas tanto dentro como fuera de la lingüística variacionista. Mientras que algunos difieren en las interpretaciones de los resultados, la mayoría de estudios de esta variable comparten un conjunto de suposiciones clave sobre la naturaleza de la variación involucrada. Este trabajo aborda algunos de estos

^{*} Grateful thanks to John Coleman, Lesley Milroy, Kirk Hazen, Ricardo Bermúdez-Otero and various conference audiences for their invaluable input to this paper. And to Sali Tagliamonte for stimulating my interest in (t,d), and for supplying the raw data.

^{*} New College, Oxford OX1 3BN England.

supuestos, demostrando cómo un examen fonético detallado de los datos plantea problemas fundamentales que sugieren que, en ausencia de evidencia independiente contraria, (t,d) se modela mejor como un proceso de habla conectada, aunque sea cognitivamente determinada.

Palabras clave

caída de -t,d, variación, fonética, fonología, proceso de habla conectada

1. Introduction

The deletion of word-final /t/ and /d/ in English consonant clusters (e.g. *find* > [faind] ~ [fain]) was first identified as a (socio)linguistic variable in African American English (AAE) (e.g., Labov *et al.* 1968), but it has subsequently been studied in a wide range of varieties by variationists and, more recently, theoretical phonologists. Most studies of this variable since the 1970s build on the foundational work of Guy (e.g., 1977, 1991) and they almost all share a number of assumptions:

- I. (t,d)¹ is an abstract, phonological variable rule, which applies to words ending in /t,d/-final clusters;
- II. (t,d) applies in both the lexical and post-lexical phonology;
- III. (t,d) governs discrete, binary alternation between surface [t,d] and zero (at least in the lexical phonology), which can be detected reliably by careful auditory analysis;
- IV. the alternation is conditioned overwhelmingly by (a) the segmental phonological context, (b) the morphological structure of the word (more deletion occurs in monomorphemic words such as *tent* than in morphologically complex words like *kept* or *sailed*) and (c) (in more recent studies) lexical frequency effects;
- V. the rule is ubiquitous in English world-wide.

¹ The abbreviation '(t,d)' is used here for convenience to represent both the linguistic variable whose variants are the presence or absence of [t,d], and the rule which variably deletes the consonant.

I to III have been found to hold consistently for AAE, where high levels of deletion are found (albeit differentially) before both consonants and vowels, and deletion is found word-internally as well as word-finally (cf., e.g., Green 2002). These phenomena are found in other varieties, for example, in North Carolina (Butters 1989); in Bermudian and Jamaican English we find examples such as (1) - (2) and (3) respectively:²

- (1) and go out to the pond [po[®]n] and pick the kite sticks
- (2) and he does pasting [peisin] and I like I said
- (3) you have to look at the date on the product $[p_{JJ}d_{k}h]$

However, in other varieties there are much more substantial quantitative differences between preconsonantal and prevocalic rates of deletion, and word-internal deletion is not found, leading Labov (1972: 9) to conclude that the two types of variety have different rules with respect to cluster reduction, one phonological and other "low-level" phonetic.

Assumption V has nevertheless remained in the literature and it has more recently been called into question further by studies which failed to find the expected effects of morphological structure in some varieties of English (e.g., Tagliamonte & Temple 2005, and Renwick *et al.* 2014, for British English; Hazen 2011, for Appalachia). The present study was prompted by these findings, and by methodological difficulties encountered in the analysis of the data for Tagliamonte & Temple (henceforth "T&T"). Its aim is to address a question posed by Wolfram about (t,d) over twenty years ago: "Is it simply enough to note whether the cluster is reduced or not, or must one note finer phonetic points of detail in terms of the cluster?" (Wolfram 1993: 211). The detailed analysis focuses primarily on the assumptions listed as III and IV(a) above, highlighting some limitations of the widely used analytical method of auditory-impressionistic transcription and exploring their implications for those assumptions and, by extension, for I and II. I thus ask not only what are the "finer phonetic points of detail" but why they matter for analyses of (t,d) as a variable phonological rule and what alternative

² Thanks to Rosemary Hall (p.c., 2014) for these examples from her data.

R. TEMPLE

analyses they suggest. The data are taken from the set analysed by T&T, and come from the York corpus of British English³ described in Tagliamonte (1998).

Section 2 presents an examination of the phonetic details of a sample of (t,d) tokens, focusing particularly on those with following obstruents or nasals, which are reported in the literature to have the highest rates of deletion (e.g., 55% in T&T: 293; Table 4). This shows how III is difficult to maintain in a high proportion of cases because it is not possible to detect unambiguous evidence of the deletion of the word-final /t,d/. In Section 3 a further sample of tokens demonstrates that IV(a) is equally problematic, since determining the strictly sequential context of application of a deletion rule is confounded by other phonetic/phonological processes and by the non-sequentiality of some phonetic cues. §4 addresses the implications of these observations for the analysis of (t,d) and considers whether it might not better be characterised as a Connected Speech Process (CSP) than an abstract phonological rule.

2. The interpretation of phonetic data

On the face of it, (t,d) is a relatively straightforward variable to model, involving categorical alternation between the absence and a surface phonetic realisation of an underlying word-final stop. It is generally acknowledged that a coronal stop following a token constitutes a "neutralizing environment" (Guy 1980: 4) and tokens in such contexts are routinely excluded from analyses on the grounds that it is not possible to tell whether a stop produced in that context is just a reflex of the following consonant or a reflex of both that and the word-final stop. However, the phonetic analysis and coding of the data for T&T showed that such difficulties arise in far more cases than merely the tokens which are conventionally excluded on the grounds of neutralisation. They are not the first to be aware of such problems. Wolfram's question, quoted above, is followed by a warning: "Although some analysts have extracted data by simply counting the consonant cluster as overt or not, I think this is an unwise move, since it presumes that

³ Collected with the support of the Economic and Social Research Council of the United Kingdom (ESRC) under Research Grant #R000238287.

all the relevant linguistic categories potentially affecting the incidence of the variable have been determined" (Wolfram 1993: 211). However, it is difficult to find evidence in the literature that his caution has been heeded. The interpretational problems T&T encountered with the raw data are grouped here somewhat arbitrarily; other groupings and other labels are possible, and the problems illustrated for each group overlap, sometimes to a considerable degree. They all concern phenomena which are instantly recognisable as normal to phoneticians familiar with CSPs (e.g., Farnetani 1999) and which have been much studied since the early invention of such articulatory techniques as palatography. This section will first review what constitutes neutralisation and then examine some further phenomena which can make it equally difficult to determine whether deletion has or has not applied. The working assumption is that if (t,d) is a phonological rule any observed phonetic reflex of underlying /t,d/ must mean that the rule has not applied, and any ambiguities in the phonetics must call into question whether or not it has applied.

2.1 Neutralisation

As already mentioned, the so-called "neutralising" environment is a context where problems in identifying variants have long been acknowledged: "... in word-final consonant clusters it is necessary to exclude clusters which are immediately followed by a homorganic stop (e.g. test day) from the tabulation since it is sometimes impossible to determine whether the final consonant of the cluster is present or absent" (Wolfram 1969: 48). The exclusion of "neutralisation" contexts seems to have been normal practice since Wolfram's study, although there appears to be no in-depth discussion of exactly which contexts should be excluded for this reason. Only one of the studies consulted by T&T (Bayley 1995) also excludes tokens with following interdental fricatives, on the grounds that they are frequently realised as stops by Bayley's Tejano subjects.

However, there are other following consonants which could arguably also have this kind of neutralising effect on the variation, but which appear never to be mentioned. The most notable is [n], which is also articulated with apical/laminal occlusion at the teeth/alveolar ridge. It might be argued that the presence of nasality would always differentiate the following nasal from the oral coronal stop, and stops, particularly voiceless ones, are often clearly audible even if there is no release before the following nasal. Where this is not the case, though, it is not clear how the analyst is to determine whether or not the speaker intends a coronal gesture for the following /n/ to pertain also to the /t/ or /d/ that precedes it. Moreover, nasality as a phonetic property is notoriously non-segmental, that is it is rarely strictly co-temporal with all the other properties of the segment to which it "belongs". In (4), for example, the [s] is followed by a brief, nasalised puff of aspiration and a partially devoiced nasal consonant:

(4) they try their <u>best</u> not [bɛsʰ nɒʔ] to stay on⁴

As with /t#d/ and other accepted "neutralisation" sequences, release of the wordfinal plosive would in this token not be expected in normal casual, unscripted speech. The nasality is clearly audible from the end of the [s], but it is very difficult to determine whether [h[~]] is actually a reflex of an underlying /t/ with nasal assimilation or nasal release, or whether the /t/ has been deleted and the nasal is merely partially devoiced. These decisions cannot be made on an *ad hoc* basis: decisions of principle need to be taken as to what is to be deemed a sufficient cue to the surface realisation of /t/ or /d/. Discussions of such principles tend, when they occur in the literature, to be limited to consideration of whether segmental variants such as flaps or glottal stops count as deletion, whereas (4) illustrates a context where the question is what subsegmental properties are sufficient to constitute absence of deletion, in this case whether the voicelessness is ascribable to the /t/ or to the juxtaposition of /s/ and /n/ alone.

Moreover, the problem is not limited to determining whether a surface reflex of /t,d/ is present: a case can also be made for treating other following consonants sharing alveolar or dental articulation with /t,d/ as neutralisation contexts on the grounds that in some sequences it is not at all clear that [t] or [d] on the one hand and zero on the other are both likely pronunciations. For example, in /st#s/ sequences in certain

⁴ All numbered examples from (4) onwards are taken from T&T's data. Phonetic judgements throughout the paper are based on a combination of visual inspection of spectrograms / waveforms and auditory analysis.

syntactic / discourse contexts (e.g. "at the <u>last</u> second"), where one might ask whether [sts] is truly a normal pronunciation in natural, rapid speech. Indeed, none of the sixteen tokens of /st#s/ in the present data set was pronounced with any surface reflex of /t/. Such problems are, however, not limited to potential "neutralisation" contexts, and we now turn to examine some areas which, I would argue, also need principled decisions about how to interpret the data and which in some cases are impossible to interpret definitively with only auditory and acoustic information.

2.2 Masking effects

The term "masking" is used here to denote the possibility of an articulatory gesture, possibly an incomplete one, which is physiologically and/or acoustically concealed by the articulation of surrounding consonants.⁵ Where there is a following vowel, the duration of the stop closure, the audible release and the visible formant transitions into the vowel make a surface reflex of the (t,d) token easy to identify, as in (5) and (6):

- (5) *er Simon and I kept in touch* [k^hɛp̈t^hıntʊtʃ]
- (6) *if if a project or* [piəʊdʒɛʔtʰɔ·] *contract comes up*

Figure 1 is a spectrogram of part of (6) showing the preceding /k/ realised as a glottal, a clear closure period and a release showing formant transitions consistent with an alveolar plosive reflex of the word-final /t/ of *project*.

⁵ For the sake of conciseness, a broad definition of masking is adopted here whereby gestures need not be anterior to the coronal gesture (as they are in the examples from Browman & Goldstein (1990) discussed below), since the acoustic consequences of the latter can also be masked by an overlapping velar closure, which would prevent the build-up of intra-oral pressure necessary to produce a coronal release burst.



Figure 1. Spectrogram of "<u>project</u> or" (6); male speaker.



Figure 2. Spectrogram of "<u>kept</u> me occupied" (7); male speaker.

In the absence of a release, by contrast, whether or not deletion of word-final /t,d/ has occurred is much less certain, as is the case with (7), illustrated in Figure 2:

(7) having this lego <u>kept</u> me [k^hɛ̯pmi^j] occupied for years.

As Figure 2 shows, there is glottalisation of the vowel of *kept* and possibly glottal reinforcement of the [p], but auditory analysis reveals that there is also unambiguous sustained bilabial closure. The following [m] is clearly visible. There is no evidence in the spectrogram or auditorily of a [t] between the [p] and the [m], but it is not possible to state categorically whether there is or is not an apical gesture present. ⁶ This unreleased /p/-to-homorganic /m/ sequence is, of course, exactly what one would expect from a native speaker of British English in connected speech (cf. e.g., Cruttenden 2008; Nolan 1992). Even assuming the absence of a lingual gesture, the presence of glottalisation might be interpreted as a reflex of /t/ in a glottal stop, but this interpretation is no more straightforward: the presence of a masked glottal stop is no easier to identify, and the creaky voicing on the preceding vowel and in the diphthong of the following word, clearly apparent in Figure 2, means that this could just be a function of the speaker's register.

Many tokens showed this kind of effect. In (7) the place of articulation of the

⁶ The relatively short duration of the closure in *kept* compared to the /p/ of *occupied* is ascribable to a rapid deceleration of speech rate and cannot necessarily be taken as an indication of /t/ deletion.

preceding and following consonant is the same, but (8) and (9) demonstrate that this is not necessary for masking to occur:

(8) well it was all pressed bits of [pissbiftsə] meat you know

(9) but there was all <u>old</u> carpets [<code>ɔlkhaph1?s</code>] and pictures.

In each case there is a preceding apical gesture towards the alveolar ridge. Since word-final stops are not obligatorily accompanied by audible release (and arguably not normally so in this type of context), the absence of an audible or visible release burst cannot be taken as unambiguous evidence for deletion of /t,d/: in (8) the blade and tip of the tongue could have raised from their fricative position to form a closure during the articulation of the "following" [b], just as the side(s) of the tongue could have raised to complete a post-lateral closure in (9). In both cases, any coronal release would be auditorily masked by the closure of the following stop. It is, of course, equally possible that the tongue tip/blade was never raised for [k] closure in (9), but it is impossible to tell either way without direct articulatory data.

Masking is particularly problematic where there is glottalisation of the preceding consonant and with combinations of preceding nasals and following plosives or nasals. (10) is taken from the same clause as (6):

(10) if if a project or contract comes [kontua?-komz] up.

Again, the preceding and following segments are unproblematic: there is a clear closure into a glottal reflex of the preceding /k/ of *contract* and a clear velar release of the initial plosive of *comes*. Again it is not possible to state categorically that there is not an apical [t] gesture present, but if this were the case the glottal gesture would have to be released before the release of a [t] and crucially before the velar closure for the following /k/, in order for the presence of the /t/ to be perceived independently or to show up on the spectrogram. Alternatively, given that a glottal stop is a common reflex of /t/, this could be construed as a further neutralising context since the presence of a

"preceding" glottal stop makes it impossible to detect whether a /t/-specific glottal reflex is present or not, or to decide whether the glottal is a reflex of /k/ or /t/ or both — see below for further discussion.

The parallel problem with preceding nasals is illustrated in (11) and (12):

(11) you know we were educated, <u>trained</u> people [tie·np^hi·pl] / [tie·ndp^hi·pl]

(12) they were over a thousand quid $[\theta a^{\circ} z \eta kwid]$ each

Occasionally, such cases could be disambiguated from spectrographic evidence, for example, a sharp cessation and resumption of voicing with word-final /t/ followed by a voiced stop, ⁷ but unsurprisingly, the majority are more like (11), represented spectrographically in Figure 3. The energy showing faintly between the [n] and the [p] release here is from the interviewer speaking over the informant; the closure period for /p/ is unambiguously voiceless. Prior to that it is possible to see the nasal energy falling off in frequency, but there is no stretch of non-nasalised voicing consistent with a fully voiced [d]. The lack of voicing could be explained by the word-final assimilatory devoicing characteristic of many Yorkshire speakers, but in the absence of a release this potential explanation is of no help in determining whether or not the word-final stop is present.

Tokens with following nasals or plosives rarely have released [t,d], and those which do have audible release usually involve hesitation or a prosodic pattern signalling a pragmatic or discourse effect. This is the case in (13) and Figure 4, where the speaker is introducing the computer game Minesweeper as the source of his friend's problems with distraction at work and produces a micro pause after *found* followed by a lengthened diphthong in the first syllable of *Minesweeper*:

(13) and he <u>found</u> Minesweeper [faʊndd maːınswiːpʰə] have you played Minesweeper?

⁷ Nine /d/ tokens with preceding /n/ and following voiceless consonants were devoiced and so also identifiable in this way (total number of /ndC/ = 72). The picture for /t/ is complicated by the fact that the majority of preconsonantal /nt/ tokens were glottalised (24/31), the proportion rising to 14/16 with following stops/nasals.

Examples (11) and (13) were produced by different speakers and the durations are different, but the spectral pattern in *found* (Figure 4) is almost identical, *mutatis mutandis*, to that in *trained* (Figure 3): in each case there is clear formant structure throughout the voiced portion of the closure for [n(d)] and no voicing bar extending beyond the end of the formants, as there would be in a canonical voiced [d]. The plosive release in (13) is completely voiceless, though not aspirated. This is again quite normal in English and it is difficult to see on what grounds one could state definitively whether or not the stop in (11) (Figure 3) has been deleted (to do so, a combination of palatography and nasal aerometry would be necessary, which is impractical for naturalistic studies). It is thus hard to see the justification for extrapolating from these and the other examples in this section a phonological rule of deletion, which would also apply to, say, the final /t/ of 'l've never seen the film *Gorillas in the <u>Mist</u>* [mis(t)]', as opposed to a CSP which would not apply in the latter case.⁸



Figure 3. Spectrogram of "<u>trained</u> people" (11); female speaker.



Figure 4. Spectrogram of "<u>found</u> mines[weeper]" (13); male speaker

2.3 Assimilation

The problem of masking is compounded in cases of assimilation across the (t,d) token. Again, this is particularly an issue with nasals, which frequently assimilate to the

⁸ An invented example is given here, since there is not a single example of a sentence-final coronal stop cluster with deletion in the York dataset, a fact which would not be coincidental under a CSP analysis.

place of articulation of a consonant following (t,d). When the underlying token is voiceless, it is sometimes possible still to detect a glottalised reflex of it, as in (14):

(14) *she's on a <u>different</u> plane* [difim_[?]ple-n].

Potentially masked reflexes of /d/ are, however, much harder to detect, as in (15), where the speaker is describing an early record player, and (16):

(15) a a a <u>sound</u> box [saumboks] was only a diaphragm

(16) we built, um, Bradford <u>combined</u> court [khəmbaïŋkhɔː?] centre.

2.4 Data interpretation: overview

The phenomenon of masking, with or without assimilation, might seem to pose purely practical problems, and an argument could be adduced from the point of view of perception that the masking causes the hearer not to hear a reflex of /t,d/ and it is thus reasonable to model its perceived absence as a result of deletion. However, the generally accepted treatment of "neutralisation" in (t,d) by excluding tokens in (following) neutralising contexts, on the grounds that it is impossible to perceive whether the (t,d) token is deleted or not, demonstrates that (t,d) is evidently modelled on the basis of production rather than perception. Since masking and neutralisation introduce the same uncertainty in the first step of the analysis, that is deciding whether a token is realised or not, they should at the very least be treated in the same way: either neutralised tokens should be included in the analysis because they form part of what the hearer hears (and presumably recognises as (t,d) sites), or potentially masked tokens should be excluded because, as with neutralisation, it is impossible for the analyst or the hearer to detect whether deletion has occurred.⁹ Given that production and perception must ultimately be linked, this decision might still be construed as merely an operational one, but it must nevertheless be addressed and it cannot be given

⁹ A decision to exclude all these tokens would of course severely curtail the analysable data set, rendering it in fact impracticable.

proper consideration without also considering the abstract model of the behaviour of (t,d). The implications of such difficulties will therefore be taken up again below. However, CSPs also pose problems concerning the interaction of any abstract (t,d) rule with other phonological / phonetic processes and we shall first examine these.

3. The relationship between (t,d) and other phonetic/phonological processes

Assumption IV(a) is predicated on the occurrence of /t,d/ in sequence between the preceding and following segments. These are assumed in the literature to be the underlying adjacent segments, yet there is surface evidence in many tokens of other phonetic/phonological processes. In this section I address firstly the issue of how these other processes might interact with (t,d) and then whether a strictly sequential analysis of the phonological constraints on (t,d) is possible at all in some cases.

3.1 Interaction with other phonological processes

It could be argued that the assimilation cases presented above constitute evidence in support of a rule of word-final coronal stop deletion: the assimilation in (16) would thus be argued only to occur because the /d/ between the nasal of *combined* and the velar plosive of *court* has been deleted before the postlexical rule of assimilation across the word boundary applies. However, in (14) assimilation of the /n/ in *different* to the place of articulation of /p/ in *plane* occurs across the glottal reflex of the word-final stop, showing that segmental adjacency is not a prerequisite for assimilation. By extension, there might well be an inaudible apical gesture in (15) and (16), but again it is impossible to tell.

Masking of the variable is not the only problem posed by assimilation for the analysis of (t,d): assimilation also causes difficulties with determining what the phonological context is when the rule applies. Thus in (14) above we might ask whether the preceding context is a syllabic [m] or a coda /n/ prior to undergoing assimilation to the following /p/.



Figure 5. Spectrogram of "contact sports" (17); male speaker.

The most statistically robust effects of phonological context have concerned manner rather than place of articulation. However, there are other phonological processes interacting with (t,d) that do affect manner of articulation, and even the major class membership of the preceding and following context. Again, we shall demonstrate the problem using individual tokens.

In (17) there is a clear release of the [t], so the token is an unambiguous example of non-application of the deletion rule:

(17) he was a bit wet when it comes to <u>contact</u> sports [k^hpnt^ha ?t spo:?s]

The following segmental context is unproblematically [s]. However, the preceding context is less straightforward: /k/ is realised as a glottal, which raises the question of what exactly the segment was when the rule applied, [k] or [?]. It might be argued that what matters for the rule is that [?] is a stop, and its place of articulation is not important, but phonetically it is realised as creak on the /a/ vowel, as shown in Figure 5, and thus in a way which is qualitatively very different from [k]. Of the 169 preceding stops in the York data, 71 are phonetically full glottal stops and five are glottalised; glottals in total thus represent nearly 7% of the data set and 45% of preceding stops (76/169),¹⁰ so this is a far from trivial question.

 $^{^{10}}$ All but one are preceding /k/s, so glottals account for 69% of preceding /k/ (75/109). The other token is /p/.

A similar problem is posed by vocalised /l/, as in (18):

(18) So she told me off [theomisf] for shouting at her

There are ten such tokens in the data set and one, (19), where there is no obvious sequential reflex of /I/ at all:

(19) my friend told me right [thəmuai] yesterday

In these and other cases of the absence of a preceding phonetic consonant, the question arises of how long in the derivation the underlying cluster remained a cluster and so subject to the (t,d) rule. Whereas tokens with preceding phonetic laterals have a deletion rate of 19% (total N=104), of the ten tokens¹¹ where the word-final consonant is preceded by a vowel in the surface form, six (60%) have the final consonant deleted, so /l/-vocalisation *would* appear to have an effect on deletion.

Questions concerning the ordering of rules also affect the following phonological context. In tokens like (20), where the /t/ coalesces with the following /j/, the same question arises: what is the following context when the (t,d) rule applies, in this case postlexically?

(20) like [the baby] <u>kept</u> you up $[k^{h}\epsilon p] t^{j} \sigma p]$ 24 hours a night

Following /h/ is particularly problematic in this respect. In (21) the following context is phonetically a vowel,¹² but underlyingly it is consonantal. What, then, is the following context when the rule applies?

(21) we was walking down Micklegate and we grabbed him [guabdim]

¹¹ The problem would, of course, be more serious in other varieties of British English where /l/vocalisation is more common. Note that Walker (2012) excludes tokens with preceding /l/ because of the prevalence of vocalisation in his Toronto data.

¹² 39 of the 62 pre-/h/ tokens (63%) are followed by a phonetic vowel.

The rate of deletion in the 62 tokens with following /h/ is actually just 11%, only marginally higher than the 8% deletion rate for following vowels suggesting that even without /h/-dropping, the classification of /h/ with other following obstruents and nasals is erroneous.

All these problems are compounded when the processes affecting adjacent consonants also affect (t,d), as illustrated in (22), where glottalisation might be applying to /k/ and/or /t/, and the order of application of glottalisation and/or (t,d) is impossible to determine:

(22) ... I w- worked part-time [wä?pä?taim] in funerals

We return to this issue in the following section.

3.2 Segmental sequentiality

Examples (7) and (22) above raise a further question, albeit one which is partly bound up with the other issues discussed in this section, that is the possibility that a phonetic reflex of (t,d) might not occur sequentially between its "preceding" and "following" segments. (7) is reproduced here for convenience:

(7) having this lego <u>kept</u> me [k^hɛ̯pmi^j] occupied for years.

The spectrogram of the token in Figure 2 above shows the audible glottalisation on the vowel of *kept* and into the [p] closure. It is well known that the phonetic cues to segmental identity are not restricted to the temporal slot implied by linear, segmental representations. The cueing of coda voicing by the duration of the preceding vowel is a commonplace, for example. So it might be argued that there *is* a reflex of /t/ present in the *kept* of (7), although it is not sequentially aligned in the word-final position. Again, this is a topic which merits further experimental exploration (into both perception and production) beyond the scope of the present paper, but again even on the present evidence the problem is raised of whether such tokens merit being classified as having undergone deletion, when reflexes arguably persist on adjacent segments.

In (7), there is clear oral articulation of the unreleased bilabial [p] of *kept* as well as the glottalisation. By contrast, voiceless velar stops immediately followed by another stop in York English (and many British varieties) are frequently realised as glottals without any velar articulation.¹³ These tokens pose a different problem for classifying segments in sequence: in (23) the [t^h] of *worked* is released so [?] and [t^h] can be taken as sequential reflexes of /k/ and /t/ respectively:

(23) and that was where my dad worked and [w3.?than] where the Barbican...

However, this is not possible in (22), or (24) and (25), which are all from different speakers:

(24) She <u>knocked</u> straight [np?stuɛi] into us yeah

(25) being an infant teacher was helpful in that <u>respect</u> because [JISbg?·bikoz].

The preceding segment in each case is realised as a glottal stop, and it appears that the (t,d) token is absent. A parallel example, (10), was discussed under *Masking* above, but even if there were no masked alveolar gesture, [?] is also a possible pronunciation of (t,d) in this variety, as evidenced in (26), so it remains impossible to disambiguate whether [?] in (24) and (25) is a reflex of /k/ or /t/ or both.

(26) you felt as [fɛl?əz] if you moved you'd fall off

It would be necessary to carry out detailed phonetic comparisons of a number of tokens with potential sequences of glottals to establish whether there is, for example, a regular pattern of variation between a lengthened [?] in *worked* versus a shorter glottal reflex of /k/ in (*I*) *work*, which would indicate (although not conclusively) that there was an undeleted /t/ in this token of *worked*.¹⁴

¹³ Preceding /p/ may also be realised as a glottal stop, as in <u>except</u> us [i?sɛ̯?'t^hʊs], but this is rare.

¹⁴ For a *caveat* on such interpretations of duration see Kühnert & Hoole (2004).

The questions raised in §3 cannot be dismissed by arguing that the rule relates to abstract phonological units or categories of sonority, major class features etc.: in order to carry out variable rule analysis, the analyst has determine for each token what the preceding context is, and it is crucial to know what that context is at the point when the variable rule applies. This is particularly important in cases where the preceding context could be a vowel, which means the cluster may not actually be a consonant cluster when the rule applies, and equally so where the following context may be a consonant or a vowel, given that following consonant versus following vowel has been found to have the most robust statistical effect on (t,d) since the very earliest studies. With an iterative lexical-phonological rule, such problems are intractable. It is impossible to determine whether the chicken of rule application came before or after the egg of, say, /l/-vocalisation.

4. Discussion

The survey of problems in this paper is intended to be illustrative and is not exhaustive, but even these affect a substantial proportion of the preconsonantal tokens, and so pose a challenge to Assumption III. The number of tokens potentially affected by each phenomenon, together with the number coded by T&T as deleted, is given in Table 1, which shows that they amount to 26% (83/325) of all tokens with following obstruents or nasals, that is the group which is consistently found to favour deletion most in the literature. Of these potentially problematic tokens, 83% were coded, rightly or wrongly, as deleted, a much higher proportion than for following obstruents as a whole (55%). Moreover, if the 62 pre-/h/ tokens are set aside on the grounds that they have probably been misclassified by being grouped with other obstruents, the proportion of problematic preconsonantal cases rises to 32% (83/263). Thus for almost one third of the tokens in the set with the highest rate of apparent deletion it is in fact difficult either to be certain that deletion has actually occurred or to know what the phonological context is when the rule applies or, indeed, both; and this is not an exhaustive tally. Thus, although this investigation began as an exploration of the methodological problems encountered during the analysis of the data for T&T, the data

Example	Sequence	Number of tokens	Number coded as deleted	
§2.1	/st#s/	16	16	
(4)	bilabial-to-bilabial masking	4	3	
(5)	/s/-to-bilabial masking	12	10	
(7),(20),(22), (23)	glottal-to-C masking/ glottal	15	10	
(8),(9),(10),(13),(14)	ambiguity /ndC ^[stop] / ¹⁵	31	27	
	TOTAL:	83	69	(83%)
Total following obstruents/nasals (inc. /h/):		325	179	(55%)
(excl. /h/):		263	172	(65%)

reviewed raise more than simply methodological issues.

Table 1. Numbers of tokens in problematic contexts including following obstruents / nasals.

4.1 (t,d) as a variable rule

Although variable rules have their roots in transformational generative phonology, their ontological status has been a matter of debate (see, e.g., Fasold 1991; Mendoza-Denton, Hay & Jannedy 2003; Walker 2012): do they represent a convenient statistical tool for modelling variation or are they a model of speakers' competence? Whatever the general answer to this question, the linguistic characterisation of (t,d) espoused in the literature entails that the rule be a phonological rule proper, at least so far as morphological class and preceding context are concerned, that is, it applies in the lexical phonology (as well as post-lexically). The question thus arises of how this particular rule fits into the phonology as a whole. It is unproblematic for lexical processes strictly associated with the derivation of verbal forms, such as the deletion (or epenthesis) of the suffix vowel of {-ed} and voicing agreement of the final consonant, to occur before the variable deletion rule applies. However, the indeterminacy of the ordering of the rule with respect to processes affecting preceding and following consonantal segments, illustrated in §3, clearly does have direct bearing on any phonological analysis. As we have seen, indeterminacy also surrounds whether the rule has even applied in many cases, suggesting that any re-evaluation of (t,d) must go beyond addressing the ordering of rules and instead re-examine the nature of the rule itself, including whether it is in

¹⁵ Including tokens assimilated to following place of articulation.

fact a phonological rule at all.

Guy (1977) argues that while the effects of following segments might be explained in low-level phonetic terms, the effect of following pause and, moreover, the crossdialectal differences in its ranking with respect to other constraints, may not. Further, the consistent differences between deletion patterns in words of different morphological structure also necessitates a phonological model. Guy concludes, therefore, that (t,d) is "a case where phonological variation cannot be accounted for by the sort of 'general functional conditions' suggested by Kiparsky [(1972)], but rather probably must be considered 'a rule of grammar'" (Guy 1977: 9) in all varieties, not just AAE. Thereafter (t,d) has routinely been treated as a phonological rule and, as mentioned above, it has been studied in the light of various phonological phenomena: core syllabification, the OCP and, most famously, as a rule of Lexical Phonology. The LP account is generally taken as given, and (t,d) has been used to support some quite fundamental theoretical claims, for example, Coetzee & Pater (2011) or Bermúdez-Otero (2010), so the questions raised above have potentially far-reaching implications.

The problem posed by phonetic gradience in the realisation of (t,d) is in fact noted, and discussed in some detail, in an unpublished paper by Myers (1996). As he points out, whereas postlexical processes might be gradient, lexical rules are generally held not to be (cf., e.g., Kiparsky 1985). The evidence for gradience in (t,d) thus poses a problem for a rule which is crucially both lexical and postlexical and is presented in categorical terms in the literature. Kiparsky allows for the possibility of rules being both categorical (lexically and postlexically) and gradient (postlexically), but the problem remains of how to determine empirically what is a categorical and what a noncategorical application of (t,d). Myers develops a distributional method of distinguishing between the two based on the duration of the consonantal portion of the acoustic waveform from the offset of the pre-cluster vowel to the onset of the following word. This is not the forum to debate the advantages and disadvantages of Myers' methodology. Rather, we may ask why it might be deemed necessary to go to the trouble of developing such methods for (t,d) in the face of the indeterminacy (with currently available methods of investigating natural continuous speech) of categoricity versus gradience due to phenomena such as masking.

148

The answer for Myers, and currently the only linguistic answer to this question, lies in the interaction of (t,d) with morphology: "One aspect of the dilemma that will arise seems unshakable: Guy's evidence that in certain dialects of American English t-deletion is both lexical and postlexical. Specifically, Guy has shown that t-deletion interacts with morphology in such a way that it must be analysed as applying both within the lexicon as well as in a domain larger than the word" (Myers 1996: 5). Independently, Bermúdez-Otero (2010) also addresses the implications of the partial gradience of (t,d), and he also argues for a two-step derivation because of its morphological sensitivity. However, as already noted (§1), several studies have cast doubt on the robustness of a statistical effect of morphology on (t,d) variability consistent with the Lexical Phonology model of (t,d).¹⁶ Moreover, allowing for partial gradience in the realisation of the word-final stop itself still leaves unsolved the problems of rule-ordering with respect to other processes and of sequentiality, which were examined in §3. Thus the abstract model of (t,d) leaves many basic questions unanswered.

4.2 (t,d) as a Connected Speech Process

The phenomena affecting the analysis of (t,d) are mostly common CSPs, in British English at least, and viewed thus, they occur precisely in the contexts where one would expect most tokens to be perceived as having the final consonant deleted. One model of phonology which does claim to be able to integrate such complex phonetic observations is Articulatory Phonology (AP), and (t,d) and masking more generally feature prominently in Browman & Goldstein's (1990) early paper. Figure 6 shows an acoustic waveform and the trajectories followed by pellets on the major articulators during the pronunciation of *nabbed most*. The figure illustrates yet more starkly the need for more abstract models than AP to address the question of whether (t,d) is a rule of production or perception: the acoustic output here would clearly count perceptually as deletion, and yet there is a very clear production gesture of the tongue blade corresponding with

¹⁶ Temple (in prep.) examines this issue in more detail.

an underlying /d/.

Most pertinent to the question of the nature of (t,d) in a model other than AP are the similarities Browman & Goldstein (1990) observe between cases such as Figure 6 and cases of variable assimilation across word boundaries not involving word-final clusters, as illustrated in Figure 7. Here the final alveolar nasal of *seven* assimilates to the following stop in *plus* but again an alveolar gesture remains. Wherever the phenomena described by Articulatory Phonology belong in a linguistic model, it is apparent that these two cases are very similar. In the case of *seven plus*, a non-AP approach would recognise the (variable) assimilation fairly uncontroversially as an albeit regular gradient phonetic CSP of English; given the similarities between that and the case of *nabbed most*, it would seem that the onus is on those wishing to espouse a more abstract phonological model of (t,d) to demonstrate that the two cases are sufficiently different to justify the treatment of the latter as the result of a different, categorical phonological rule. Temple (2014) provides further, detailed evidence of the extensive parallels between the phonetic behaviour of (t,d) consonants and that of other wordfinal stops in connected speech, both singletons and in clusters.¹⁷





Figure 7. X-ray pellet trajectories for "seven plus seven" [sɛvmִ#plʌs] (Browman & Goldstein 1990, p. 22, Figure 11b)

It does not necessarily follow from this that the variability in the realisation of (t,d)

¹⁷ There have been many more recent experimental phonetic studies and some acoustically informed variationist studies with a direct or indirect bearing on word-final (t,d), for example, Mitterer & Ernestus (2006), Raymond *et al.* (2006), Schuppler *et al.* (2009). However, there have to my knowledge been no published studies focussing specifically on (t,d). This in itself is telling. Some of the studies and their implications for an alternative analysis of (t,d) are discussed further in Temple (2014).

consonants should simply be put down to automatic effects of physiological constraints on connected speech. Even without espousing the phonological model of AP, the burgeoning field of sociophonetics¹⁸ is testament in itself to the fact that phonetic variability which is not part of the core phonology is not necessarily an automatic result of Kiparsky's "general functional conditions impinging on speech performance" rendering it, "unnecessary to investigate variation of this type" (Guy 1977: 4). Since phonetic detail may be controlled in structured ways by speakers (see, e.g., Docherty et al. 1997; Carter 2003; Stuart-Smith & Scobbie 2014), there is no a priori reason why (t,d) should not be a variable phonetic phenomenon¹⁹ which behaves in a structured manner. All the data in Temple (2014) are taken from the same dialect, but there is evidence of speakers systematically varying the characteristics of (t,d) and other word-final stops in different discourse contexts. This and the existence of cross-dialectal variation in deletion rates before pauses is not sufficient to justify an abstract phonological account of (t,d): it could simply be that more general CSPs are deployed differently in different dialects. Again, positive evidence would be needed to demonstrate that the pre-pausal differences are specific to (t,d) and so warrant an abstract rule.

Moreover, in a CSP account of (t,d) issues of rule ordering and sequentiality cease to be problematic: coalescence (as in (20) above) and the co-occurrence of CSPs such as lenition/elision of word-final /t,d/ and /l/-vocalisation (as in (18)) are to be expected (cf. Nolan 1992); indeed, Temple (2014) provides examples of co-occurrence of both lenition and fortition of (t,d) with lenition and fortition phenomena beyond the word containing the cluster. Viewing (t,d) as a particular instance of more general CSPs would seem, then, to provide the most satisfactory solution to the multiple problems surrounding the abstract model. The evidence for cognitive control of the variation nevertheless necessitates a means of formally modelling it. One possible model is AP, but that would only account for a subset of the phenomena observed here, and

¹⁸ Sociophoneticians would see this as an umbrella term rather than a 'field', covering a broad range of approaches which share a common commitment to exploring interfaces between phonetic and variationist theory and practice, rather than representing a monolithic school of thought, in the same way as Laboratory Phonology is an umbrella term indicating a commitment to the use of empirical methods in phonology.

¹⁹ This paper takes an agnostic stance regarding whether a dividing line might be drawn between postlexical phonological processes and phonetic CSPs. The crucial issue here is that (t,d) is not characterisable as a categorical (rather than gradient) phonological rule.

moreover, Lichtmann (2010, p.c.) has demonstrated with articulatory data that not all cases of "deletion" involve gestural overlap, which is the driving mechanism of that model. Once again Guy suggests a possible new solution: Guy *et al.* (2008) come to the data with a very different approach from the phonetically driven one espoused here, but on the basis of lexical and frequency effects found in their study of (t,d) in New Zealand English, they suggest a role for the "highly enriched lexical representations" of exemplar-based models. Such models are, of course, equally suitable for capturing the complex phonetic richness of representations in context.

5. Conclusions

This paper has demonstrated how detailed phonetic analysis of word-final consonant clusters ending in /t,d/ calls into question the key assumptions underpinning their modelling in terms of an abstract variable phonological rule applying to the wordfinal segment. That analysis depends on two key assumptions: that there is reliably detectable binary alternation between a surface consonant and zero (Assumption III) and that it is possible reliably to determine the preceding and following segments (IVa). Well known connected-speech phenomena such as assimilation and articulatory masking pose problems for both these assumptions which go beyond the methodological issue of determining when deletion has occurred and in what phonological context. The evidence of gradience and similarities in behaviour between (t,d) consonants and other CSPs noted by, e.g., Browman & Goldstein (1990), Myers (1996) and Temple (2014) suggest the need to revisit the widely held assumption that there is an abstract variable rule of /t,d/-deletion in all varieties of English which applies in both the lexical and post-lexical phonology (I, II, V). The failure of several studies to find a robust statistical effect of morphological class on the variability (IVb) and the fact cross-dialectal differences can be accounted for equally well in terms of CSPs remove the two main independent arguments for an abstract phonological analysis. It would appear, then, that Wolfram's caution regarding ignoring fine phonetic detail was wellfounded, as was Labov's distinction between AAE-like and other varieties with respect to this phenomenon (see §1 above): when (t,d) is investigated "from the starting point of combinatorial phonetics", as suggested in the conclusion of T&T, it looks very much like a post-lexical CSP.

References

- BAYLEY, R. (1995) "Consonant cluster reduction in Chicano English", Language Variation and Change, 6, 303-326.
- BERMÚDEZ-OTERO, R. (2010). "Currently available data on English t/d-deletion fail to refute the classical modular feedforward architecture of phonology", paper presented at the 18th Manchester Phonology Meeting. http://www.bermudez-otero.com/18mfm.pdf>
- BROWMAN, C. & L. GOLDSTEIN (1990) "Tiers in articulatory phonology, with some implications for casual speech", in J. Kingston & M.E. Beckman (eds.), Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech, Cambridge: Cambridge University Press, 341-376.
- BUTTERS, R. (1989) The Death of Black English: Divergence and Convergence in Black and White Vernaculars, Berlin Germany: Peter Lang.
- CARTER, P. (2003) "Extrinsic phonetic interpretation: spectral variation in English liquids", in J.K.
 Local, R.A. Ogden & R.A.M. Temple (eds.), *Phonetic Interpretation. Papers in Laboratory Phonology VI*, Cambridge: Cambridge University Press, 237-252.
- COETZEE, A. & J. PATER (2011) "The place of variation in phonological theory", in J.A. Goldsmith, J. Riggle & A. Yu. (eds.), *The Handbook of Phonological Theory*, 2nd ed., Oxford: Blackwell, 401-434.
- CRUTTENDEN, A. (2008) Gimson's Pronunciation of English, London: Hodder.
- DOCHERTY, G.J., P. FOULKES, J. MILROY, L. MILROY & D. Walshaw (1997) "Descriptive adequacy in phonology: a variationist perspective", *Journal of Linguistics*, 33, 275-310.
- FARNETANI, E. (1999) "Coarticulation and connected speech processes", in W.J. Hardcastle & J. Laver (eds.), *The Handbook of Phonetic Sciences*, Oxford: Blackwell, 371-404.
- FASOLD, R. (1991) "The quiet demise of variable rules", American Speech, 66, 3-21.
- GREEN, L.J. (2002) African American English. A Linguistic Introduction, Cambridge: Cambridge University Press.

- GUY, G. (1977) "A new look at -t, -d deletion", in R.W. Fasold & R.W. Shuy (eds.), *Studies in Linguistic Variation*, Washington DC: Georgetown University Press, 1-11.
- GUY, G. (1980) "Variation in the group and the individual: the case of final stop deletion', in W. Labov (ed.), *Locating Language in Time and Space*, New York: Academic Press, 1-36.
- GUY, G. (1991) "Explanation in variable phonology: an exponential model of morphological constraints", *Language Variation and Change*, 3, 1-22.
- GUY, G., J. HAY & A. WALKER (2008) "Phonological, lexical and frequency factors in coronal stop deletion in early New Zealand English", Abstract for paper at Laboratory Phonology 11, Wellington, New Zealand. http://labphon.org/LabPhon11/publish/abstracts.php>
- HAZEN, K. (2011) "Flying high above the social radar: coronal stop deletion in modern Appalachia", *Language Variation and Change*, 23, 105-137.
- KIPARSKY, P. (1972) "Explanation in phonology", in S. Peters (ed.), *Goals of Linguistic Theory*, Englewood Cliffs, NJ: Prentice Hall, 189-227.
- KIPARSKY, P. (1985) "Some consequences of Lexical Phonology", Phonology Yearbook, 2, 85-138.
- KÜHNERT, B. & P. HOOLE (2004) "Speaker-specific kinematic properties of alveolar reductions in English and German", *Clinical Linguistics and Phonetics*, 18, 559-575.
- LABOV, W. (1972) Language in the Inner City. Studies in the Black English Vernacular, Philadelphia: University of Pennsylvania Press.
- LABOV, W., P. COHEN, C. ROBINS & J. LEWIS (1968) A Study of the Nonstandard English of Black and Puerto Rican Speakers in New York City, Washington DC: US Office of Education.
- LICHTMAN, Karen (2010) "Testing Articulatory Phonology: variation in gestures for coda /t/", *Illinois Language and Linguistic Society Conference 2 (ILLS2)*, University of Illinois at Urbana-Champaign.
- MENDOZA-DENTON, N., J. HAY & S. JANNEDY (2003) "Probabilistic sociolinguistics: beyond variable rules", in R. Bod, J. Hay & S. Jannedy (eds.), *Probabilistic Linguistics*, Cambridge MA: MIT Press, 97-138.
- MITTERER, H. & M. ERNESTUS (2006) "Listeners recover /t/s that speakers reduce: evidence from /t/-lenition in Dutch", *Journal of Phonetics*, 34, 73-103.
- MYERS, J. (1996) "The categorical and variable phonology of variable t-deletion in English", paper presented at the International Workshop on Language Variation and Linguistic Theory, University of Nijmegen, September 1995. http://www.ccunix.ccu.edu.tw/~Ingproc/tdeletion-manuscript.pdf>

- NOLAN, F. (1992) "The descriptive role of segments: evidence from assimilation", in G.J. Docherty & D.R. Ladd (eds.), *Papers in Laboratory Phonology II*, Cambridge: Cambridge University Press, 261-280.
- RAYMOND, W.D., R. DAUTRICOURT & E. HUME (2006) "Word-internal /t,d/ deletion in spontaneous speech: modeling the effects of extra-linguistic, lexical and phonological factors", *Language Variation and Change*, 18, 55-97.
- RENWICK, M.E., R.A.M. TEMPLE, L. BAGHAI-RAVARY & J.S. COLEMAN (2014) "Deletions in Big Data? The phonetics of word-final (t,d) in the Audio BNC", Paper presented at the biennial colloquium of the British Association of Academic Phoneticians, Oxford, 9th April, 2014.
- SCHUPPLER, B., W. VAN DOMMELEN, J. KOREMAN & M. ERNESTUS (2009) "Word-final [t] deletion: an analysis on the segmental and sub-segmental level", in *Proceedings of the 10th International Conference of the International Speech Communication Association (Interspeech 2009)*, 2275-2278.
- STUART-SMITH, J. & J. SCOBBIE (2014) "Derhotacisation in Scottish English: a sociophonetic journey", in C. Celata & S. Calamai (eds.), Advances in Sociophonetics, Amsterdam: John Benjamins, 59-96.
- TAGLIAMONTE, S. A. (1998) "Was/were variation across the generations: View from the city of York", Language Variation and Change, 10, 153-191.
- TAGLIAMONTE, S.A. & R.A.M. Temple (2005) "New perspectives on an ol' variable: (t,d) in British English", Language Variation and Change, 17, 281-302.
- TEMPLE, R.A.M. (2014) "Where and what is (t,d)? A case study in taking a step back in order to advance sociophonetics", in C. Celata & S. Calamai (eds.), *Advances in Sociophonetics*, Amsterdam: John Benjamins, 97-134.
- TEMPLE, R.A.M. (in prep.) "Towards a phonetically accountable model of (t,d)."
- WALKER, J.A. (2012) "Form, frequency and function in phonological variation", *Language Variation and Change*, 24, 397-415.
- WOLFRAM, W. (1969) A Sociolinguistic Description of Detroit Negro Speech, Washington DC: Center for Applied Linguistics.
- WOLFRAM, W. (1993) "Identifying and interpreting variables", in D. Preston (ed.), American Dialect Research, Amsterdam: John Benjamins, 193-221.