

University of Groningen

Laryngeal contrast and phonetic voicing

Jansen, Wouter

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2004

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Jansen, W. (2004). *Laryngeal contrast and phonetic voicing: A laboratory phonology approach to English, Hungarian, and Dutch*. s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 5

Experiment 1: regressive voicing assimilation in English

In chapter 4 I described three phonetic characteristics that a voicing assimilation rule would need to be classified as coarticulation-based: (1) since coarticulatory voicing assimilation is likely to be driven mainly by the articulatory gestures involved in the production of voicing distinctions it should be triggered by actively (de)voiced obstruents only; (2) the process should be reflected only by the voicing of target obstruents, and those features that are mechanically linked to voicing, such as frication duration; (3) the process does not categorically erase phonetic distinctions between tense and lax target obstruents (cf. 11 in 4.1.2).

The aim of this chapter is to test the three main predictions of a coarticulation-based view of RVA across word boundaries by means of an acoustic investigation of regressive assimilation in English. Although English is often regarded as a language with little or no voicing assimilation at word boundaries, it does in fact allow for all three predictions to be tested because English possesses both actively voiced and actively devoiced sounds, and because it maintains a contrast between fortis and lenis obstruents in word-final contexts.

The two experimental results reported below broadly support the hypothesis formulated in 4.4 above which holds that RVA at word boundaries is an articulatory process: Actively voiced English /z/, and to a lesser extent, actively devoiced /t, s/ all appear to trigger some form of RVA, in contrast to passively voiced /d/. The phonetic reflexes of this process are mostly limited to the phonetic voicing of the target obstruent, and consequently the process is non-neutralising.

5.1 Predictions

One of the principal predictions of the phonetic theory of RVA described in 4.1.2 is that the capacity of a sound for triggering assimilation is a function of its voicing target: actively devoiced sounds should trigger devoicing, actively voiced sounds are expected to increase the voicing of a preceding sound if possible, whilst passively voiced sounds (sounds lacking a voicing target) should not affect the voicing of a preceding sound. The specific predictions that can be derived from this theory for aspirating varieties of English are listed in (15). Fortis obstruents /p, t, k, tʃ, f, s, ʃ/ are expected to cause some degree of devoicing in a preceding obstruent, because they are likely to be actively devoiced. Given that the lax fricatives /v, z, ʒ/ are actively voiced (see section 2.3), these sounds should cause an increase in the duration of the voicing interval of a preceding obstruent. The lax plosives /b, d, g, dʒ/ on the other hand, should act as an intermediate, ‘neutral’ environment for a preceding obstruent similar to that provided by a following sonorant. Both the sonorants (cf. 2.1) and the word-initial lax plosives of aspirating varieties of English (2.2.1) are arguably passively voiced, which means that they lack articulatory targets gestures related to the production of voicing distinctions, and consequently they are unable to pass on such gestures to neighbouring sounds by means of coarticulation.

- (15) Predictions of a coarticulation-based approach to voicing assimilation regarding obstruent sequences in aspirating varieties of English
- a. English obstruents fall into 3 classes in terms of their influence on (the voicing of) a preceding obstruent. (1) fortis obstruents trigger devoicing; (2) lenis *fricatives* cause an increase in the voicing of a preceding obstruent; (3) *lenis* stops behave as an intermediate, ‘neutral’ category, as do sonorants. Cf. (11c)
 - b. The assimilatory effects of fortis obstruents and lenis fricatives are limited to voicing and features mechanically dependent on the production of voicing distinctions. Cf. (11a)
 - c. English RVA at word boundaries is a gradient process that is not neutralising in most instances Cf. (11b)

In addition, it follows from the mechanism underpinning RVA in the phonetic theory, first, that the effects of tense obstruents and lax fricatives on a preceding obstruent should be limited to voicing and those phonetic features mechanically dependent on the production of voicing distinctions. Second, the phonetic theory predicts that RVA across word boundaries in English is a gradient process which is non-neutralising on most occasions, even for the phonetic features that it does affect.

5.2 Previous descriptions of the phonetics of English obstruent clusters 123

By contrast, under the strict interpretation of a lexical feature analysis assumed in 4.1.1, differences in voicing between lax plosives and fricatives are predicted to have no impact on the occurrence of RVA: either the process applies in a manner-symmetric fashion, or it does not apply at all before lax obstruents (see prediction 10c in 4.1.1). Second, even if a lexical feature model can *represent* manner-asymmetric RVA at word boundaries it still predicts that where it applies, the process affects all phonetic cues to [tense] (10a). Third, it predicts that, as a result, even manner-asymmetric RVA will always act in a phonetically neutralising fashion.

The acoustic investigation reported below was specifically designed to test these two contrasting sets of three predictions. The results of this investigation indicate that the behaviour of English velar stop + /t, d, s, z, r/ sequences is as predicted by the phonetic theory on all three counts in (15), and therefore warrant a revision of traditional descriptions of voicing assimilation in English. These traditional descriptions, and some of the relevant experimental literature is reviewed in the next section.

5.2 Previous descriptions of the phonetics of English obstruent clusters

As noted at several points in the previous chapters, impressionistic accounts usually describe standard varieties of English as possessing little or no regressive voicing assimilation across word boundaries. Indeed, Jones (1956) warns native speakers of French and Dutch against making the mistake of applying RVA in English obstruent clusters (e.g., Jones's §851). This point is echoed by Gimson (1994), who does, however, claim that at the boundaries of 'close-knit' groups of words, lenis fricatives (but normally not lenis stops) may devoice completely when preceding a fortis obstruent. Moreover, according to Gimson, this form of RVA may also shorten the preceding vowel, although this phenomenon is judged to be "relatively rare" (Gimson 1994:257). This description of RVA to English fortis fricatives is reminiscent of the account of a regressive devoicing rule found in Yorkshire dialects of English provided by Wells (1982a). According to Wells's description, this rule is triggered by all fortis obstruents and affects voicing as well as durational correlates of [tense].

Instances of regressive assimilation to lax obstruents documented in impressionistic accounts invariably involve function words or word internal contexts. One instance is the voiced pronunciation of plosive + alveolar fricative clusters in words containing orthographic <x> such as <exam>, [ɪgzæm]; <exhibit>, [ɪgzɪbɪt]; <excerpt> [ɛgzɜ:pt]. These clusters presumably originate from an older (and invariant) form [ks] which was subsequently affected by the process of [s]-voicing that is also responsible for the pronunciation of e.g., <desire> as

[dɪzɑɪə]. According to [Borowsky \(2000\)](#) the (historical) voicing of word-medial prestress /s/ in turn triggers voiced realisations of the preceding /k/. The same author discusses a second example of apparent word internal RVA in English: the optional voiced realisation of the final alveolar fricative of prefixal <dis> before lax stops, as in [dɪzɡ aɪz] for <disguise>. This observation seems to directly contravene a phonetic approach to RVA, which predicts that English lax stops should be unable to act as triggers of the process.

[Borowsky \(2000\)](#) grants that the voiced realisations of orthographic <x> and the alveolar fricative of <dis> are optional, but her descriptions nevertheless fail to do justice to nature of voicing patterns in English medial obstruents and obstruent clusters. A number of additional observations cast doubt on her claim that the optional voicing of the medial fricative in <disguise> is due to the same (synchronic) assimilatory mechanism that underlies the (also optional) voicing of the medial /s/ in Dutch /mɪs/ + /da:d/, [mɪzda:t], *crime*. For example, judging by pronunciation dictionaries such as [Windsor Lewis \(1972\)](#), [Jones \(1977\)](#), and [Wells \(2000\)](#), the English process idiosyncratically affects the final sibilant of <dis> before /g/ and /d/ (e.g., <disdain>) but not before /b/ or /dʒ/ (cf. <disbar, disband, disjoin, disjunct>), although this apparent place of articulation effect may be an artefact of morphological transparency and/or lexical frequency. Furthermore, the final sibilant of <mis-> in e.g., <misguided>, <misgiving> is normally realised as [s] rather than [z]. Dutch RVA by contrast, is not lexically selective in this way.

More importantly, the correct generalisation about the voicing of English orthographic <s> seems to be that it can occur before sonorants as well as lax obstruents, but not before tense obstruents. Thus, the examples in (7c) in the previous chapter match <Osborne>, <Osgood>, <Marsden>, <Neasden>, which all have [z] rather than [s], but e.g., <Oscar, osprey> are normally pronounced with [s]. Similarly, as illustrated in (16), postpausal and postsonorant weak forms of <is> can be realised with [z] before lax obstruents and sonorants, but not before tense obstruents ([Lakoff, 1972](#); [Selkirk, 1980](#)).

(16) Voicing of weak <is> in English (examples from [Selkirk 1980](#))

Orthography	Phonetic form
<Is Jack going?>	[(ɪ)z dʒæk ɡoʊɪŋ]
<Is Will going?>	[(ɪ)z wɪl ɡoʊɪŋ]
<Is Pete going?>	[(ɪ)sp ^h ɪt ɡoʊɪŋ]

As argued in chapter 4, observations that (neutralised or weakened) obstruents are voiced before both lax obstruents and sonorants do not necessarily imply that either or both of these classes actively contribute to the effect. Moreover, such observations are consistent with a coarticulation-based approach to RVA as described in 4.1.2 to the extent that the voicing process is motivated independently (e.g., by passive voicing). Given that English alveolar sibilant voicing

5.2 Previous descriptions of the phonetics of English obstruent clusters 125

is indeed motivated independently, an articulation-driven account of its origins would say that the process applies freely before passively voiced lax obstruents and (equally passively voiced) sonorants because they cannot influence the voicing of a preceding sound (cf. prediction 15a), but that it is blocked by assimilation to actively devoiced fortis obstruents, which can. In other words, the observation that words such as <Osborne> and <disguise> are commonly produced with [z] does not necessarily constitute evidence against a phonetic view of regressive voicing assimilation in English.

The picture of English voicing assimilation I have drawn so far is broadly speaking reflected in the generative literature on the topic, which tends to classify English as a language without RVA across word boundaries and sometimes as a language in which only the laryngeal specification of fortis obstruents is visible to the phonology (Harris 1994; cf. chapter 8). Perhaps in part because of this picture, instrumental investigation of English obstruent clusters with mixed underlying [tense] specification has been limited, and in all but one case that I am aware of, does not allow for the specific predictions of the phonetic theory of RVA to be tested against those of a lexical feature analysis. However, the picture emerging from the single study in question is considerably more encouraging for the phonetic theory than the one sketched by impressionistic accounts.

The quantitative (acoustic) study of laryngeal contrast and voicing in American English fricatives conducted by Stevens et al. (1992) shows a clear effect of following context on the voicing of fortis and lenis fricatives. For example, in their corpus lenis fricatives (/v, z/) have on average 29 ms of voicing preceding a fortis fricative (/f, s/), which increases to 58 ms before another lenis fricative or a vowel. However, since Stevens et al. (1992) do not provide separate means for fricative + vowel sequences and homogeneous (tense + tense and lax + lax) clusters, there is no baseline measure to determine whether the shorter voicing intervals in the fortis environments are the result from some form of active devoicing or whether the greater amount of voicing in the lenis contexts is the result of RVA to the lenis fricatives, or both. Furthermore, Stevens et al. (1992) do not provide tests of the statistical significance of the differences in the mean voicing values they observe.

Although statistical tests are provided in an acoustic investigation of English velar obstruents in various contexts by Myers (2002), it does not allow testing of the phonetic theory of RVA either. Whilst Myers's test stimuli contain both /z/ and /g/ as following obstruents, his (statistical) analysis does not distinguish between these two environments. Consequently, it is impossible to determine whether the slight increase in voicing he observes before lax obstruents is due to a symmetric effect of /g/ and /z/ or an asymmetric effect of /z/.

However, an early and all but forgotten study by N. Thorsen (1971) (and one that I have only become aware of after most of the work reported below

Table 5.1: Voicing duration (ms) and ratio of the closed phases of unstressed English /t/ and /d/ followed by a C₂ in the onset of a stressed syllable, as reported by N. Thorsen (1971).

C ₁ C ₂	C ₁ voicing		C ₁ + C ₂	C ₁ voicing	
	Duration	Ratio		Duration	Ratio
/tl/	35	0.53	/dl/	56	1.00
/tk/	31	0.66	/dk/	44	0.86
/tg/	41	0.87	/dg/	51	0.91
/ts/	33	0.60	/ds/	50	0.79
/tz/	64	0.82	/dz/	62	0.96

had been completed) does shed considerable light on the issues raised by the phonetic theory. This study investigates voicing and other phonetic features of [tense] in English alveolar stop C₁ + consonant C₂ sequences straddling word and morpheme boundaries in three different prosodic contexts, and crucially reports measurements for stop, fricative, and sonorant C₂s separately. Some of the mean values reported by Thorsen are represented in table 5.1.

Although the match between this voicing data and the three term classification in (15a) is imperfect, it is striking how the absolute durations of the voiced intervals of /t/ and /d/ are longer before /z/ than before /g/. Moreover, the mean C₁ voicing value for /d/ + /l/ is more or less intermediate between those for /d/ + /z/ on the one hand and /d/ + /s/ and /d/ + /k/ on the other, although /d/ + /g/ would be expected to group with /d/ + /l/ rather than with /d/ + /s/. Absolute voicing duration in /t/ is not as predicted in (15a) to the degree that /l/ appears to pattern with fortis obstruent C₂s rather than with /g/. In addition, /t/ is relatively short before /g/, and consequently its *voicing ratio* (duration of the voiced interval divided by overall duration of the closed phase) is higher there than before /z/. But note that voicing ratio is only a good indicator of ‘degree of assimilation’ if both overall duration and voicing duration generally pattern in the fashion predicted by a lexical feature analysis.¹

In addition to this (limited) evidence for manner-asymmetric RVA in English, N. Thorsen (1971) provides evidence in support of (15b) and especially (15c). No effects of assimilation are discernible in the durations of the closure phases of /t/ and /d/, whilst the length of the vowels preceding the clusters in table 5.1 clearly preserves the contrast between /t/ and /d/ (all differences significant at $p < .01$). Note however, that there is a 13 ms difference between the vowels preceding /t/ + /s/ (32 ms) and /t/ + /z/ (45 ms) in the direction

¹Interestingly, of the means provided in table 5.1 the only pairwise difference(s) that N. Thorsen (1971) lists as statistically significant at the $p < .05$ level is the difference in C₁ voicing duration and/or ratio between /d/ + /s/ and /d/ + /z/ (the type of test is not specified).

predicted by a lexical feature account.

5.3 Methods

Subjects Subjects were 4 native speakers (2 male, 2 female) of British English aged between 24 and 35, and living in or near to London at the time of recording. None of the subjects had a history of speech or hearing impairment. They were not paid for their participation in the experiment. 3 subjects, K6, L7 (both female) and R10 (male) were speakers of a south-eastern variety of British English, while the speech of the remaining subject J11 (male), displayed some characteristics of his native Lincolnshire although it had no strong local features. All 4 subjects were non-rhotic.

Materials The stimuli for this experiment consisted of consonant clusters combining a /k, g, ŋ/ C₁ and a /t, d, s, z, r/ C₂. Velar stop C₁ were preceded by a long central mid open vowel [ɜ:] (V₁)², while /ŋ/ followed low back rounded /ɒ/ (V₂). C₂ was always followed by a vowel.

The main reason to use velar rather than alveolar C₁ was that word-final /t/ is often realised as a glottal stop in British English. A different place of articulation was chosen for the C₁ consonants for segmentation purposes; the choice for velar stops over labial ones was determined by the desire to control for the preceding vowel. The choice to use alveolar C₂s was made partially because of the exceptional behaviour of lenis labiodental fricatives with regard to RVA in a number of languages (e.g., Hungarian, Russian), and partially because some claims about the phonetic basis for the nature of laryngeal contrast in fricatives have been made with specific reference to sibilants (Balise & Diehl, 1994).

- (17) English sample stimuli
- a. How does patchwork duvet translate?
 - b. How does headstrong zealot translate?
 - c. How does Hamburg dairy translate?

The clusters were located at the internal boundary of noun + noun (N₁ + N₂) constructions and further embedded within a carrier phrase (*How does _ translate?*) designed to attract nuclear stress on the second noun. Both N₁s and N₂s were disyllabic with an initial lexical stress. Thus, the rhythmic structure of the stimuli and nuclear accent placement were controlled to maximise the potential effect of RVA, which has been shown to depend on lexical stress in Dutch (see

²This vowel is transcribed in square brackets in order to side-step questions about the underlying representation of orthographic <V + r + C> sequences, as in, e.g., <work>. Note that all subjects realised such sequences with a long vowel rather than [V + ɹ].

Slis 1986 and chapter 4). Given the sparsity of English words beginning with /z/ no attempt was made to control for the lexical frequency of the target words N_1 and N_2 . For each of the 15 different consonant clusters 4 stimuli were prepared. Some sample stimuli are given in 17, with target consonant clusters in a slanted font.³

Stimuli containing the sonorant consonants /ŋ, r/ (realised by all subjects as [ŋ, ɹ] in word-final and word-initial contexts respectively) were included to create baseline conditions for the comparison of the relative effects of fortis vs. lenis C_2 on the properties of a preceding obstruent.

Procedure The stimuli were presented to the subjects in a quasi-randomised order to avoid consecutive stimuli with identical consonant clusters. The subjects were asked to repeat each stimulus three times at a comfortable rate and to read a stimulus again if they made a mistake or produced a hesitation. In total, $3 (C_1) * 5 (C_2) * 4 (\text{stimuli}) * 3 (\text{repetitions}) * 4 (\text{speakers}) = 720$ utterances were recorded.

Recordings were made onto minidisk in a sound-proofed room using a Brüel and Kjær condenser microphone (Type 4165) and measuring amplifier (Type 2609), and digitised at 22.5 kHz. Segmentation and acoustic measurements were carried out using PRAAT. 23 utterances had to be discarded because they contained small speech errors or (hesitation) pauses between C_1 and C_2 and 37 utterances were excluded because an underlying /k/ was realised as a glottal stop. In addition, all (remaining) clusters starting with a /ŋ/ are excluded from the discussion below because they are largely irrelevant to the hypotheses under consideration, leaving a total of 425 utterances for analysis.

Segmentation and measurements Segment boundaries were determined by visual inspection of waveforms and broadband spectrograms based on Fast Fourier Transforms (*FFT*) on a 5 ms Gaussian window (spectrogram bandwidth 260 Hz). The boundary between a vowel and a following plosive C_1 was placed where there was an abrupt change in the higher frequency energy, as illustrated by figure 5.1. The boundary between a C_1 plosive and a following C_2 was placed at the end of the release burst of the plosive, where *release burst* was defined as the initial transient and any following frication that could be assigned to C_1 rather than to a C_2 fricative.

59% (101 out of 171) of plosive-plosive clusters had a clear C_1 release and could therefore be internally segmented according to these criteria. In the remaining utterances where this was not the case, no boundary was placed and voicing and duration characteristics were measured for the cluster as a whole. In a few cases, mainly involving /g/ followed by /z/, the initial plosive was

³A full list of stimuli appears in appendix A.

followed by a short period of schwa-like voicing. These intervals were treated as voiced releases (analogously to the ‘embryonic vowels’ often observed after word-final lenis stops in French), and consequently their duration and voicing were assigned to C_1 . In another set of tokens the release was completely obscured by the noise of a following fricative. Here the boundary was set at the onset of the noise signal.

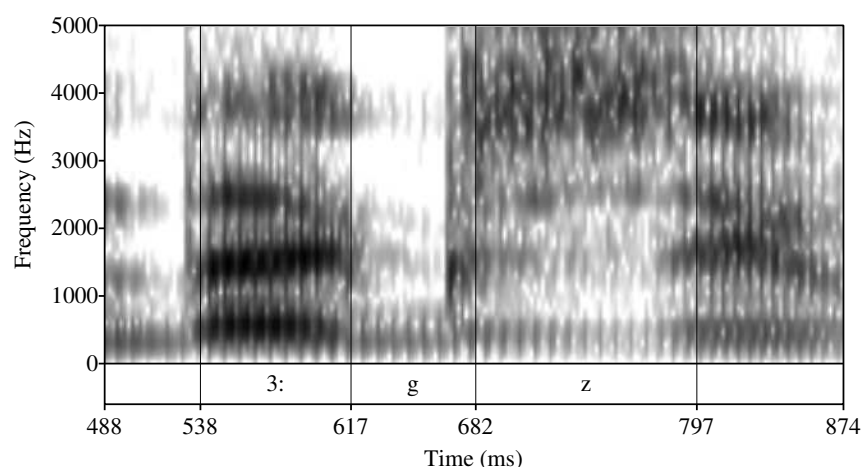


Figure 5.1: Sample spectrogram of an English /gz/ cluster. Speaker: R10 (male).

The offset of C_2 constriction was defined as the offset of frication for /s, z/ and the onset of the release burst for /t, d/. The first measurement point for F_0 was placed at 10 ms from the offset of frication for fricatives and at 10 ms after the onset of post-release voicing for plosives.

Voicing measures were determined on the basis of periodicity in the waveform and the presence of a voice bar in the spectrogram. Note that on the basis of these criteria the /gz/ cluster in figure 5.1 is voiced throughout. VOT was defined in the standard way in terms of the timing of voicing onset relative to the onset of the release burst in plosives.

The measurements that were made on the basis of the hand-segmented speech samples, as well as the relevant derived measures are listed in table 5.2, ordered by speech segment. As described in 2.2.3, preceding vowel (V_1) duration is a crosslinguistically recurrent feature of [\pm tense], which is generally considered to be particularly salient in English (Chen, 1970; Flege & Hillenbrand, 1987). F_0 and F_1 values were extracted at 10 ms intervals between 50 and 10 ms preceding the onset of C_1 closure, using the autocorrelation and Burg algorithms embedded in PRAAT 4.0. C_1 closure duration and release duration

were measured separately for two reasons. The first is that they are not necessarily part of the same cue in articulatory or perceptual terms. In theory it is therefore possible that only one of these two features turns out to cue [tense] in the subjects' speech, and in this case simply measuring overall C_1 duration might lead to a distorted picture of the reflexes of [tense] marking on C_1 and/or C_2 in terms of segmental duration. The second, more practical reason for considering C_1 closure and release duration separately is that when a stop is followed by a fricative, the release noise of the former may be partially or wholly obscured by frication noise of the latter. Again, this might distort the interpretation of C_1 overall duration. On identical grounds, C_1 voicing measures are reported separately for closure and release phases.⁴

Table 5.2: Acoustic measurements and derived measures for Experiment 1.

		Segment					
V_1		C_1		C_2		V_2	
(a)	Duration	(d)	Closure duration	(i)	Closure duration (stops)	(n)	F_0 10-50 ms after C_1 offset
(b)	F_0 50-10 ms before C_1 on-set	(e)	Release duration	(j)	Overall duration (fricatives)		
(c)	F_1 50-10 ms before C_1 on-set	(f)	Overall duration	(k)	Voicing duration (fricatives)		
		(g)	Voicing duration (2 m.)	(l)	Voicing ratio (fricatives)		
		(h)	Voicing ratio (2 m.)	(m)	VOT (stops)		

The nature of the phonetic theory of RVA demands that the phonetic features of C_2 be examined as well. It is particularly important to determine whether the subjects indeed produce tense /t/ with a long lag VOT and lax /d/ without closure voicing and a short lag VOT. Similarly, it is important to establish whether lenis /z/ has any voicing in a postobstruent environment since it would point to

⁴Only absolute voicing durations are reported in the main text of this chapter. C_1 voicing *ratios* were also calculated because they are sometimes used as a measure of RVA in other experimental studies. The interested reader can consult them in appendix A. My main motivation for focusing on absolute values of duration and voicing rather than voicing ratios is that the latter type of measure combines two acoustic features of [tense] in way that inflates the distance between two set of obstruents if both its components behave as they do in 'typical' cases of intervocalic fortis-lenis contrast. The relatively short duration and large amount of voicing of lenis obstruents both contribute to a relatively high voicing ratio, whilst the long duration and lack of voicing of fortis obstruents both contribute to a low voicing ratio. However, if either absolute duration or absolute voicing duration behaves contrary to the 'typical' pattern, the effects of underlying [\pm tense] or RVA on one feature may be (partially) canceled by the other and voicing ratio ceases to be a reliable measure.

the presence of active voicing that is critical to the predictions of the phonetic theory. The phonetic description of C_2 below also includes measurements of segmental duration and F_0 perturbations in the following vowel, but not attempt was made to determine F_1 contours, as V_2 vowel quality was not controlled for.

5.4 Results

5.4.1 Phonetic features of C_2

The data in table 5.3 and figure 5.2 indicates that the subjects use voicing distinctions to signal the distinction between tense and lax stops and fricatives as would be expected of an aspirating language. Thus, stops /t/ and /d/ can be characterised as voiceless aspirated vs. (passively) voiceless whilst the contrast between /s/ and /z/ is realised as voiceless vs. (partially) voiced. This means that prediction (15a) above is indeed applicable to the present corpus.

The mean VOTs for /t/ (70 ms) and /d/ (14 ms) fall into the standard ranges for the long lag and short lag categories, and the difference between them is highly significant according to a t-test: $t(99) = 16.18$, $p < .001$. All tokens of tense /s/ are completely voiceless, whilst /z/ has a substantial amount of voicing (77 ms). The mean voicing ratio for this obstruent is .78 (standard deviation: .22), which is fairly high in comparison with earlier studies such as Haggard (1978) or Smith (1996). Unsurprisingly, the mean difference in absolute voicing duration is statistically highly significant: $t(173) = -23.62$, $p < .001$.⁵

Table 5.3: Experiment 1: duration and voicing of C_2 . Closure duration and VOT of /t, d/, and overall duration and duration of the voiced interval for /s, z/. All values in ms, and pooled across preceding contexts (/k, g/). Standard deviations in brackets.

C_2	VOT	Closure duration	N
/t/	70 (15)	56 (16)	44
/d/	14 (19)	71 (15)	57
Voicing		Duration	N
/s/	0 (0)	132 (18)	92
/z/	77 (31)	99 (17)	83

F_0 microprosody seems to signal the distinction between tense and lax C_2 obstruents, too. Figure 5.2 plots the mean F_0 at five measurement points from 10 to 50 ms into the vowel following C_2 for the two male subjects R10 and J11. It

⁵All data on stop + stop clusters in this section pertain to sequences that could be internally segmented, unless indicated otherwise. The result is a fairly large discrepancy in the number of cases for plosive and fricative C_2 s.

shows how 10 ms into the vowel, F_0 values for /t, s/ on the one hand and /d, z, r/ on the other are roughly 20-25 Hz apart, and then gradually converge as time progresses. Both the magnitude of the F_0 differences and the grouping of lenis (passively devoiced) /d/ and (actively voiced) /z/ with sonorant (passively voiced) /r/ are in line with earlier observations in the literature. A one-way ANOVA for C_2 laryngeal specification (tense obstruents vs. lax obstruents vs. sonorant) confirms that there is a highly significant effect of the phonological status of C_2 on F_0 at the onset of a following vowel: $F(1,171) = 24.05$, $p < .001$. Tukey and Scheffe post hoc tests indicate that lax /d, z/ and /r/ are both significantly distinct from tense /t, s/ (both $p < .001$) but not from each other.⁶

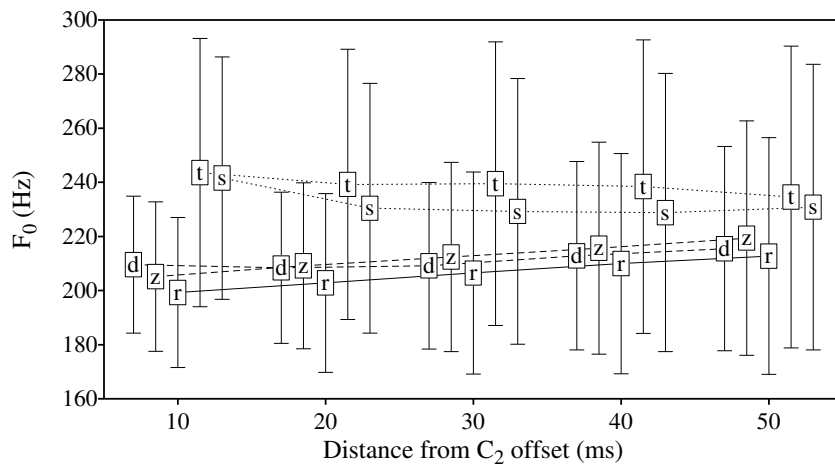


Figure 5.2: Experiment 1: F_0 (Hz) 10-50 ms into the vowel following C_2 , for female speakers only. Both internally segmented and unsegmented stop + stop clusters included. Error bars represent the mean ± 1 standard deviation.

Finally, fricative (frication) duration but not stop closure duration behave according to the typical [\pm tense] pattern. On average, /s/ is 33 ms longer than /z/, and this difference is highly significant according to a t-test, $t(137) = 12.51$, $p < .001$. The 15 ms difference in closure duration between /t/ and /d/ is also statistically significant ($t(99) = -4.94$, $p < .001$) but patterns in the ‘wrong’ direction. As closure duration is not known to be a cue to [\pm tense] in word-initial contexts this finding hardly topples any theories (cf. 2.2.3), and it is hard

⁶Utterances from the two female speakers were excluded from figure 5.2 and the ANOVA because of a considerable difference in overall F_0 level between the male and female subjects. However, the behaviour of the female subjects with regard to post- C_2 F_0 perturbations is highly similar to that of the female subjects, with a maximal difference between /d, z, r/ and /t, s/ of approximately 35 Hz.

to say whether any meaningful interpretation can be assigned to it.

5.4.2 Phonetic features of C₁ in the baseline environment

Sonorant /r/ was included as a baseline C₂ environment for the phonetic expression of the contrast between /k/ and /g/. Since /r/ is both passively voiced and phonologically unmarked for [tense] it can be treated as a ‘neutral’ context for the purposes of both the phonetic theory and lexical feature analyses of RVA.

The data summarised in figures 5.3, 5.4, 5.6 and 5.5 further below indicates that the 4 subjects use many of the phonetic features reviewed in chapter 2 to distinguish /k/ from /g/ before sonorants.⁷ For example, the bottom two bars of the top panel of figure 5.6 show how /g/ has a shorter release phase than /k/ in this environment (21 vs. 33 ms), and a marginally shorter closure stage, too (44 vs. 50 ms). Similarly, the bottom two bars of figure 5.5 show that the mean duration of vowels preceding /g/ and /k/ pattern as would be expected on the basis of the literature: on average [ɜ:] is 27 ms longer before the lax stop than before its tense counterpart (99 vs. 72 ms). Both the release duration and preceding vowel duration differences are highly significant according to t-tests: $t(77) = 4.54$, $p < .001$ and $t(77) = -5.47$, $p < .001$. The difference in C₁ closure duration is also statistically significant, but the effect is clearly less strong: $t(77) = 2.41$, $p < .02$.

The bottom panel of figure 5.4 shows that there is a difference of 21 ms in overall voicing duration between /g/ and /k/ before /r/ (43 vs. 22 ms). The difference in overall voicing ratio is .43 (.27 vs. .70: note that this difference is similar to that obtained by N. Thorsen 1971). All possible measures of C₁ voicing yield statistically highly significant differences between /k/ and /g/. For instance the t-test result for closure voicing duration is $t(77) = -5.53$, $p < .001$. It is often suggested that the duration of the preceding vowel is the primary cue to [tense] in English word-final obstruents, but these findings suggest that it is, or can be, supported by voicing distinctions.

Of the remaining components of the low frequency feature proposed by Kingston & Diehl (1994, 1995), the test subjects only appear to employ F₁ perturbation (in the present context: see further below). Figure 5.3 plots the first formant contour of the vowel [ɜ] at 10 ms intervals between 50 and 10 ms preceding the onset of C₁. The downward slope of this contour is steeper before /g/ than before /k/, which results in a 26 Hz difference (476 vs. 502 Hz) at 10 ms before the onset of C₁. This pattern agrees with data reported in the literature on the topic (cf. Stevens et al. 1992 and other references mentioned in section 2.2.3 above) and the same applies to differences in F₀ at 10 ms before the onset

⁷Exact values for the standard deviations indicated by the error bars in figures 5.4, 5.6 and 5.5 are given in appendix A.

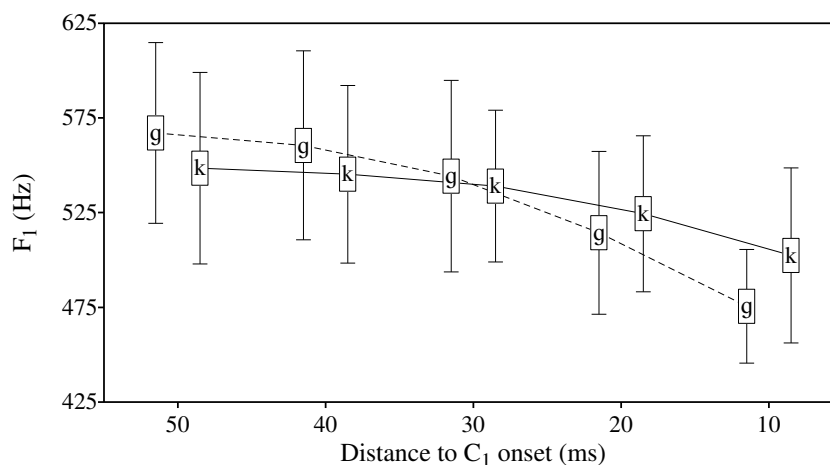


Figure 5.3: Experiment 1: F_1 (Hz) at 5 points from 50-10 ms preceding C_1 onset in /k/ + /r/ and /g/ + /r/ sequences. Error bars represent the mean ± 1 standard deviation.

of /k, g/: 192 vs. 183 Hz for the female speakers and 150 vs. 138 Hz for the two male subjects. However, only the difference in F_1 (at 10 ms) is statistically significant according to a t-test: $t(77) = 3.13$, $p < .005$.

Interestingly, correlations between the individual phonetic features discussed here are generally weak and often not statistically significant. This holds in particular for correlations between V_1 duration and the length of (parts of) C_1 . The two strongest correlations are between F_1 value at 10 ms before C_1 and V_1 duration ($r = -.35$, $p < .005$) and between the former of these and overall C_1 voicing ($r = -.32$, $p < .005$). The general absence of strong correlations between the values of the individual phonetic features of [tense] is consistent with a view in which they are traded off against each other for perceptual reasons (and manipulated independently). It is inconsistent with models that seek to reduce the cluster of correlates of [tense] to the reflexes of a single or relatively few articulatory gestures.

5.4.3 C_1 voicing in obstruent clusters

Having established the phonetic features of tense and lax velar stops in the base-line environment, it is now possible to assess whether they undergo any form of assimilation in potentially non-neutral environments. The patterning of C_1 voicing in obstruent clusters shows that this is indeed the case, and in a fashion that is entirely consistent with prediction (15a) of the phonetic theory. There is

an increase in voicing duration before /z/ but not before /d/ vis-à-vis the —/r/ baseline environment. In addition, the voicing data lends some support to prediction (15c) since voicing duration appears to partially preserve the underlying distinction between /k/ and /g/, even where RVA does apply.

Figure 5.4 provides the means for the duration of the voiced intervals of the C₁ closure and release stages as segments of bars indicating the overall voicing duration of C₁. The most striking generalisation emerging from the data in this figure is the difference in voicing between velar obstruents preceding /r/ and /d/ on the one hand, and /z/ on the other. For example, there is barely any difference in the C₁ closure (1 ms) or C₁ overall (3 ms) voicing of /kd/ and /kr/ sequences, whilst the same measures show a marked increase in voicing in /kz/ clusters. Clusters starting with /g/ behave in exactly the same fashion.

The data in figure 5.4 prompts two additional observations. First, the fortis obstruents /t/ and /s/ appear to have an assimilatory effect on the voicing of a preceding /g/. Relative to the baseline context, the length of the voiced interval of the closure stage of /g/ drops by 10 and 12 ms before /s/ and /t/ respectively. On the other hand, /t/ and /s/ have little effect on the voicing of a preceding /k/. Second, voicing duration preserves the contrast between underlying /k/ and /g/ before /d/, where no assimilation occurs, but there is a hint that even where assimilation does occur, voicing distinctions are incompletely neutralised. Thus, the average overall voicing duration of /g/ is marginally greater than that of /k/ before /t, s, z/.

A number of ANOVAs were carried out to examine whether these impressionistic observations stand up to statistical scrutiny. A first set of three-way ANOVAs for *C₂ laryngeal specification* (/t, s/ vs. /d, z/) * *C₂ manner of articulation* (/t, d/ vs. /s, z/) * *C₁ laryngeal specification* (/k/ vs. /g/) was performed on the voicing data for clusters composed of /k, g/ and /t, s, d, z/. The goal of these ANOVAs was to assess the apparent effects of regressive voicing assimilation and incomplete neutralisation in obstruent sequences. Their results indicate that RVA is indeed manner-asymmetric in the obstruent clusters produced by the test subjects, as predicted by the phonetic theory.

For example, the ANOVA on the C₁ closure voicing data shows highly significant main effects of *C₂ laryngeal specification*, $F(1,268) = 73.75$, $p < .001$; *C₂ manner of articulation*, $F(1,268) = 31.11$, $p < .001$; and *C₁ laryngeal specification*, $F(1,268) = 12.71$, $p < .001$. An ANOVA for C₁ overall voicing (duration + release) duration yields equally significant main effects, and both ANOVAs show a strong interaction between *C₂ laryngeal specification* and *C₂ manner of articulation*: $F(1,268) = 23.78$, $p < .001$ (closure voicing), and $F(1,268) = 28.81$, $p < .001$ (overall voicing). The main effects of *C₂ laryngeal specification* and *C₂ manner of articulation* support the idea that the voicing of C₁ is subject to assimilation to a following obstruent. The strong interactions between *C₂ laryn-*

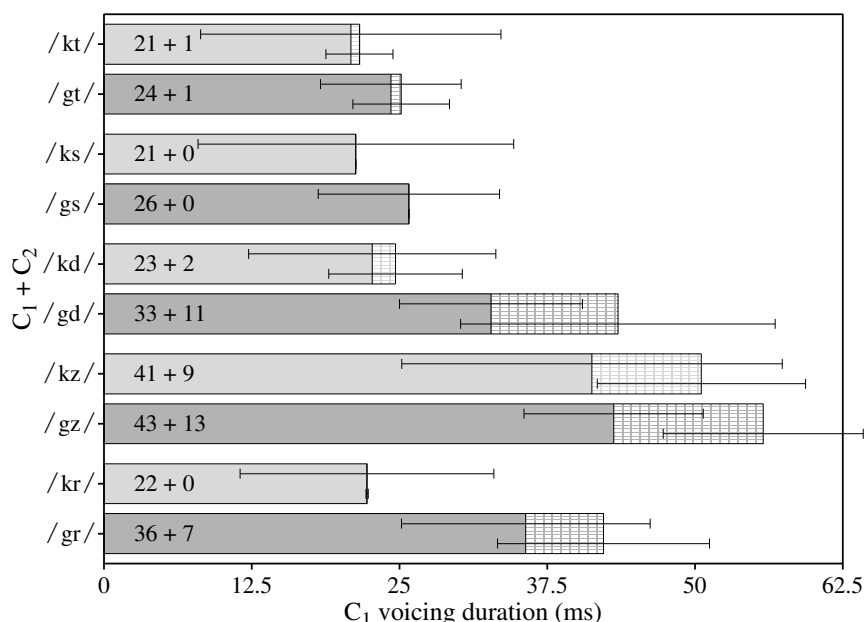


Figure 5.4: C_1 voicing duration across C_2 contexts. All measures in ms; error bars represent the mean \pm 1 standard deviation. The top panel represents mean C_1 closure and release voicing durations separately: for each bar the left-hand segment indicates the extent of voicing during the closure phase and the right hand segment represents the duration of voicing during the release stage. Exact values for these means are printed in the leftmost segment for typographical reasons. This diagram shows a marked increase in voicing, relative to the baseline environment, for both /k/ and /g/ before /z/. There is a small decrease in the voicing of /g/ when it precedes /t, s/. These observations suggest that /z/ and, to a lesser extent, the two fortis obstruents, trigger RVA. Before /d/ /k/ and /g/ behave more or less as in the baseline environment, which is an indication that the lenis plosive is unable to trigger voicing assimilation.

geal specification and C_2 *manner of articulation* indicate that both main effects are largely caused by a single laryngeal/manner class, and therefore that only 1 of the 4 obstruents under investigation is a strong trigger of assimilation. Given that it induces the greatest deviations in C_1 voicing (duration) from the baseline condition, this strong trigger is most likely to be /z/.

However, the devoicing of /g/ before /t, s/ probably also contributes something to the main effects of C_2 *laryngeal specification*. A second series of three-way ANOVAs on the closure and overall voicing data summarised in figure 5.4

with /r/ included as a separate C_2 laryngeal category ([0tense]). Tukey and Scheffe post-hoc tests on these ANOVAs show that both tense and lax C_2 environments are distinct from the baseline context provided by /r/ (as well as from each other: $p < .001$ for all pairwise comparisons). Broadly speaking, statistical analysis therefore confirms the idea that RVA is only triggered by obstruents that are actively voiced (/z/) or actively devoiced (textipa/t, s/).

The main effects of C_1 laryngeal specification finally, indicate that voicing distinctions between underlying /k/ and /g/ are not entirely neutralised. I suggested above that the principal source of this effect might be the ___/d/ context, where assimilation of voicing does not appear to occur. However, the only indication of an asymmetric preservation effect is a weak interaction of C_1 laryngeal specification and C_2 laryngeal specification revealed by the initial three-way ANOVA on the C_1 overall voicing data (obstruent clusters only): $F(1,268) = 5.17$, $p < .025$. This suggests that the 5 ms difference in C_1 overall voicing between /kz/ and /gz/ reinforces the 18 ms difference between /kd/ /gd/, and therefore that C_1 voicing distinctions between underlying /k, g/ are partially preserved before the [-tense] class as a whole.⁸

5.4.4 Segmental duration and obstruent clusters

Whereas C_1 voicing shows reflexes of manner-asymmetric regressive voicing assimilation, the same is not true of the durational measures. These phonetic features generally seem to preserve the contrast between /k/ and /g/. As a result, the segmental duration data contradicts a lexical feature analysis and largely supports prediction (15b) of the phonetic theory.

The patterning of V_1 duration offers the most unequivocal evidence for prediction (15b): the data in figure 5.6 shows only small variations of this parameter due to C_2 context. Moreover, the largest difference within a single C_1 laryngeal category in response to a change in C_2 environment occurs between /gd/ (89 ms) and /gz/ (100 ms) and can therefore not be interpreted in terms of [tense] assimilation to a following C_2 (which would predict that these two environments pattern together). The underlying contrast between /k/ and /g/ on the other hand, induces relatively large differences in V_1 duration in the expected direction (longer vowels before /g/). It seems therefore that preceding vowel duration does not assimilate, and this is confirmed by a three-way ANOVA for C_2 laryngeal specification * C_2 manner of articulation * C_1 laryngeal specification, on V_1 duration before /k, g/ + /t, d, s, z/ sequences. This ANOVA yields a significant main effect of C_1 laryngeal specification only: $F(1,268) = 72.38$, $p < .001$, and no significant interactions.⁹

⁸There were no other interactions in any of the ANOVAs reported here.

⁹Adding the unsegmented plosive + plosive clusters only strengthens the effect of C_1 laryngeal specification: $F(1,338) = 129.90$, $p < .001$ and still fails to reveal any other significant effects.

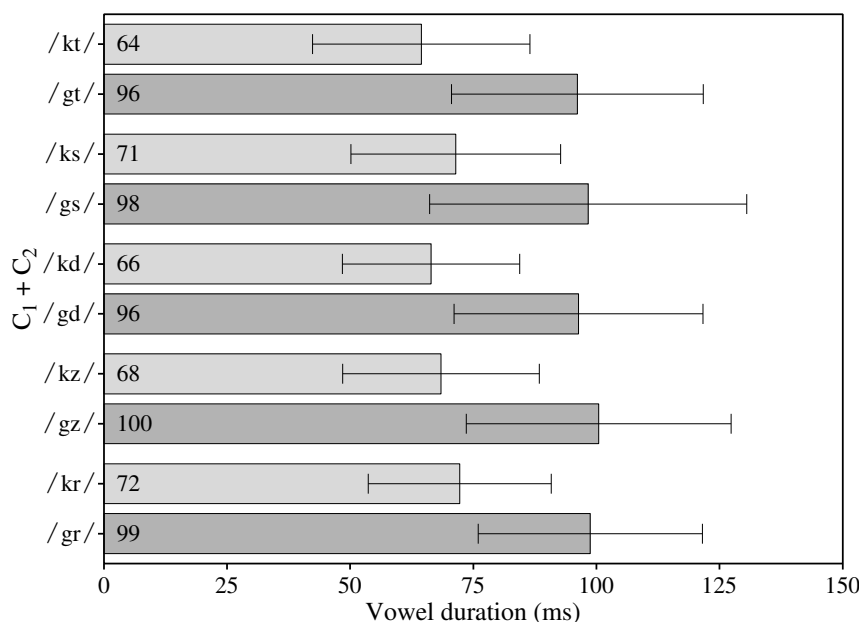


Figure 5.5: V_1 duration across C_2 contexts. All measures in ms; error bars represent the mean ± 1 standard deviation. This diagram shows that vowel duration reflects the status of the immediately following plosive (C_1) but is impervious to the value of [tense] on C_2 consonants: /k/ is preceded by a vowel of relatively short and highly similar duration across C_2 contexts whilst /g/ is preceded by longer vowels of near -identical length.

Figure 5.6 indicates that the behaviour of C_1 duration is a little more complicated. The closure interval of /k/ seems to assimilate to the [tense] value of a following obstruent to the extent that it is somewhat longer before /s/ and especially /t/ than before /d/ and /z/ respectively. However, the difference between /ks/ and /kz/ is marginal, and there does not seem to be any effect on the closure duration of /g/ that can be interpreted as evidence of [tense] assimilation. Given that /g/ does assimilate to a following /t, s/ or /z/ in terms of voicing, this results in a mismatch in the behaviour of closure duration and voicing which clearly contradicts the predictions of a lexical feature account (cf. prediction 10b in 4.1.1 above).

It is not surprising therefore that a three-way ANOVA for C_2 laryngeal specification * C_2 manner of articulation * C_1 laryngeal specification on the C_1 closure duration data (in clusters composed of obstruents only) does not reveal a significant main effect of C_2 laryngeal specification. There is a highly signifi-

cant main effect of C_1 *laryngeal specification*: $F(1, 268) = 28.07$, $p < .001$. This indicates that C_1 closure duration preserves the distinction between underlying /k/ and /g/, and inspection of the data in figure 5.6 shows that there is indeed a systematic difference in the duration of /k/ and /g/ in all but one context. In addition, the ANOVA yields a significant main effect of C_2 *manner of articulation*, $F(1,268) = 25.82$, $P < .001$, and interaction of C_2 *laryngeal specification* * C_1 *laryngeal specification*, $F(1,268)$, $p < .01$. The first of these effects is likely to stem from the relatively long closure phase of /k/ before /s/ and /z/ whilst the second probably results from the difference in the durations of /k/ closure before /t, s/ and /d, z/. Only the latter effect can be interpreted in assimilatory terms, but as I noted above, this does not vindicate a lexical feature analysis of RVA.

Finally, consider the behaviour of C_1 release duration. The first generalisation concerning this feature that emerges from figure 5.6 is that the release of /k/ and /g/ is relatively short before fricatives. As I hinted in 5.3 this is likely to be a labelling artefact caused by the overlap of release and frication noise in the acoustic signal, and it therefore seems safer to exclude cases involving a fricative C_2 from further analysis.

This leaves the sequences ending in a /d/ or /t/. The data for these clusters may seem to indicate that their internal releases are affected by some form of regressive assimilation, as on average C_1 release duration is somewhat shorter before /d/ than before /t/. However, a two-way ANOVA for C_2 *laryngeal specification* * C_1 *laryngeal specification* on the C_1 release duration data for stop + stop clusters shows that the effect of the first factor is little more than a trend: $F(1,97) = 3.29$, $p < .075$, which suggests that release duration is subject to little or no regressive assimilation. At the same time there is a weakly significant main effect of C_2 *laryngeal specification*: $F(1,97) = 6.82$, $p < .015$, which indicates that C_1 release duration at least partially preserves the underlying contrast between /k/ and /g/ (the interaction between the two factors is not significant).

5.4.5 F_0 and F_1 preceding obstruent clusters

No assimilatory patterns can be discerned in the F_0 contours preceding obstruent clusters. F_1 perturbations on the other hand, appear to show a [tense]-symmetric assimilation effect that patterns as would be expected from a rule spreading the lexical laryngeal features of C_2 obstruents: velar stops preceding /t, s/ are marked by a higher F_1 10 ms into the preceding vowel than those preceding /z, d/ (cf. table 5.4). The underlying distinction between /k/ and /g/ is erased before /t/, but in the remaining three obstruent contexts F_1 is lower before /g/ than before /k/ (as it is in the baseline environment), and this suggests that the effect of C_2 is incompletely neutralising.

A three-way ANOVA for C_2 *laryngeal specification* * C_2 *manner of artic-*

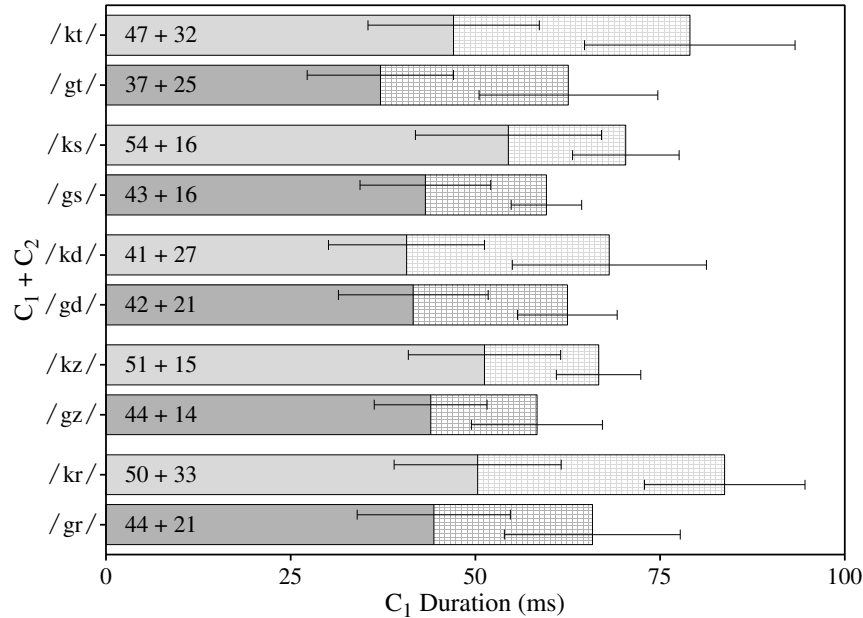


Figure 5.6: C_1 duration across C_2 contexts. All measures in ms; error bars represent the mean \pm 1 standard deviation. The diagram represents mean C_1 closure and release durations separately: for each bar the left-hand segment indicates the closure duration and the right hand segment represents the duration of the release phase. Exact values for mean C_1 closure and release duration are given in the leftmost segment for typographical reasons. This diagram provides little or no evidence for an effect of RVA on the closure duration of C_1 , because there is little or no systematic shortening (relative to the baseline context) before /t, s/ or shortening before /z/. There is some suggestion of an assimilatory effect on C_1 release duration before plosives, but this is not confirmed by statistical tests.

Table 5.4: Experiment 1: F_1 preceding obstruent clusters. F_1 (Hz) at 10 ms before the onset of C_1 . Standard deviations in brackets.

C_1C_2	F_1 at C_1 - 10 ms	N	C_1C_2	F_1 at C_1 - 10 ms	N
/kt/	508 (42)	31	/gt/	508 (34)	26
/kd/	486 (41)	26	/gd/	478 (20)	18
/ks/	519 (43)	47	/gs/	487 (28)	45
/kz/	489 (43)	36	/gz/	479 (34)	47
/kr/	502 (46)	32	/gr/	476 (30)	47

ulation * C_1 laryngeal specification on the F_1 data supports both observations. There is a highly significant main effect of C_2 laryngeal specification, $F(1,268) = 23.01$, $p < .001$, and a weaker effect of C_1 laryngeal specification, $F(1, 268) = 7.36$, $p < .01$, but no significant main effect of C_2 manner of articulation and no significant interactions.¹⁰ The absence of any effects related to C_2 manner of articulation confirms that unlike C_1 voicing, F_1 assimilates in a [tense]-symmetric fashion. As I noted in the survey of phonetic properties of the tense-lax distinction in chapter 2, lax stops appear to be marked by a lower first formant on flanking vowels regardless of their voicing targets, and so the fact that there is [tense]-symmetric assimilation of F_1 does not necessarily contradict the phonetic theory of RVA. In the light of the C_1 voicing and segmental duration data it certainly does not support a lexical feature analysis, which predicts that assimilation should be [tense]-symmetric or [tense]-asymmetric across phonetic features. However, since the articulatory underpinnings of low frequency spectral cues to $[\pm\text{tense}]$ remain unclear, the data in table 5.4 are difficult to interpret.

5.5 General discussion and conclusions

The aim of this chapter was to test three hypotheses derived from a coarticulation-based view of regressive voicing assimilation at word boundaries. The first of these hypotheses was that only actively (de)voiced obstruents should be able to trigger RVA. In chapter 4 I pointed out how a coarticulatory view of voicing assimilation correctly predicts the distribution of regressive assimilation under word sandhi within the Germanic group of languages. The data from experiment 1 indicates that English obstruents trigger RVA broadly in accordance with this view. Actively voiced /z/ and to a lesser extent, actively devoiced /t, s/ all cause deviations in the phonetic voicing of a preceding obstruent relative to a baseline sonorant context. Crucially, English /d/, which was argued in chapter 2 to be passively voiced, did not trigger any form of voicing assimilation (ignoring for the moment the effect on F_1 : see further below).

Regressive voicing assimilation to /s, z/ is qualitatively [tense]-symmetric in the sense that both lenis and fortis obstruents are able to induce changes in the voicing in at least one class of preceding obstruent. However, regressive assimilation is not always observably symmetric with regard to $\pm\text{tense}$ target sounds. For example, fortis obstruents do not affect the voicing of a preceding /k/ vis-à-vis the baseline context. The most natural interpretation of these observations is that coarticulation still applies in the relevant clusters but fails to leave a trace

¹⁰Tukey and Scheffe post-hoc tests on a second three-way ANOVA, which included the data from the baseline context shows that the [+tense] C_2 environment is distinct from both the [-tense] and sonorant environments (all $p < .001$), but that the latter two environments are not distinct from each other.

in the speech signal. For example, the devoicing gestures of a /t/ would still be anticipated during the production of a preceding /k/, but because the latter is accompanied by active devoicing measures of its own this has little effect on the voicing of the initial stop. However, this interpretation is in need of further support from articulatory data.

The second hypothesis under hypothesis was that regressive assimilation only affects the voicing of a target obstruent and those features that are mechanically linked to the production of voicing, such as frication duration in fricatives. The data gathered by the present experiment provides an almost perfect match with this hypothesis, because the behaviour of C_1 voicing but not the patterning of C_1 closure and release duration or V_1 duration can be interpreted in terms of RVA. This does not mean that C_1 duration features are not subject to modification when another obstruent follows (there are clear changes relative to the baseline context), but that these modifications cannot be attributed to the same mechanism that controls C_1 voicing. Note that the results of experiment 1 are similar to those obtained by [N. Thorsen \(1971\)](#) who does not find evidence of assimilatory effects on V_1 duration and C_1 duration characteristics either. Furthermore, the lack of regressive assimilation of C_1 duration matches Russian data presented by [Burton & Robblee \(1997\)](#).

The one remaining puzzle with regard to the results of experiment 1 is the observation that the value of F_1 towards the end of V_1 appears to be subject to manner-symmetric but tense-asymmetric assimilation to /d, z/ (assuming that — is a legitimate baseline condition for this feature). As the articulatory underpinnings of F_1 perturbations by [\pm tense] obstruents are unclear, any interpretation of this data will be speculative. Note however, that as long as the effect of obstruents on the first formant of flanking vowels can be traced to a definite articulation involved in the expression of the tense-lax contrast in English obstruents, the F_1 data does not contradict the phonetic theory (as the gesture in question would itself be subject to anticipatory coarticulation).

The third and final prediction of a coarticulatory view of regressive voicing assimilation is that the process should not be completely neutralising. This prediction is clearly borne out by the data summarised above. Even C_1 voicing, the primary feature involved in the process, tends to bear residual traces of distinctions between fortis and lenis C_1 obstruents.

Preceding vowel duration is generally regarded as the most important cue to [\pm tense] distinctions in English postvocalic obstruents (cf. [2](#)). Given that V_1 duration is entirely unaffected by any form of regressive assimilation, it seems hardly surprising that descriptions based on auditory impressions tend to regard (aspirating) English as a language with little or no RVA. However, experiment 1 indicates that this view is not entirely accurate, and that an articulation-driven form of RVA is active at word boundaries even in aspirated varieties of English.