

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*. [Thesis fully internal (DIV), University of Groningen]. 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 7

Experiment 3: voicing assimilation in Dutch three-term clusters

The aim of this chapter is to extend the comparative investigation of regressive voicing assimilation that was begun in the previous chapter in a number of ways. The experiment reported here examines the phonetic manifestation of voicing assimilation in a third language, Dutch, and in word-final clusters rather than singleton obstruents. Moreover, it assesses the influence of global variations in speaking register on regressive voicing assimilation.

The results of this experiment broadly support the predictions of a coarticulation-based view of voicing assimilation. They are also consistent with the hypothesis that Dutch dynamic final neutralisation leads to the phonetic underspecification of [tense] features, as proposed by [Ernestus \(2000\)](#). The principal conclusions of the investigation are first, that contrary to assertions in some of the literature on the topic, regressive assimilation does take place in Dutch three-term obstruent clusters with a medial fricative. Second, contrary to the pervasive assumption that Dutch RVA is asymmetrically triggered by lax plosives only, the acoustic data reported below indicate that the process is triggered by tense and lax plosives alike. Third, Dutch RVA affects the voicing of target obstruents, but not the other phonetic correlates of [tense], and therefore operates in a highly similar way to its English counterpart. Somewhat surprisingly however, there is little evidence for either an increase or a decrease in RVA at higher speaking rates.

7.1 Predictions

Dutch word-final neutralisation, regressive voicing assimilation, and postobstruent fricative devoicing were discussed in chapters 2 and 3 respectively. A brief recapitulation of the broadly accepted description of these phenomena should suffice here. According to most phonologists, Dutch has a process of word-final neutralisation that erases all distinctions between word-final tense and lax obstruents. These word-final obstruents are subject to an (optional) rule of regressive voicing assimilation when they precede a lenis plosive. Fortis plosives are generally assumed not to trigger assimilation (or to trigger it vacuously) because word-final obstruents are subject to the more general final neutralisation/devoicing process. Lenis fricatives generally do not trigger regressive assimilation but are devoiced after another obstruent. These three processes were illustrated in (4), (8), and (14) above, and repeated (in part) here as (22) for convenience.

(22) Final neutralisation and regressive voicing assimilation in Dutch

a. Final neutralisation in Dutch

UR	Plural	Citation	diminutive	Gloss
/xrap/	[χrapən]	[χrap]	[χrapjə]	joke
/krab/	[krabən]	[krap]	[krapjə]	crab
/yrat/	[χra:tən]	[χra:t]	[χra:tjə]	fishbone
/yrad/	[χra:dən]	[χra:t]	[χra:tjə]	degree

b. Regressive voicing assimilation

UR	Phonetic form	Gloss
/ve:k/ + /di:r/	[veʲ:gdi]	mollusc
/zand/ + /bank/	[zandbank]	sand bank
/vis/ + /di:fjə/	[vizdifjə]	common tern
/reiz/ + /du:l/	[reizdul]	destination

c. Progressive devoicing of lax fricatives

UR	Phonetic form	Gloss
/dræk/ + /vat/	[drəkfat]	pressure vessel
/ri:t/ + /zɑŋər/	[ritsɑŋə]	sedge warbler
/ləf/ + /zɑŋ/	[ləfzɑŋ]	eulogy
/klas/ + /vərklɛmɪŋ/	[klasvərklɛmɪŋ]	class size reduction

With the notable exception of Ernestus (2000), most accounts of Dutch laryngeal phonology subscribe to a lexical feature analysis of RVA and the fortition analysis of final neutralisation. The first part of this section summarises the (shared) predictions of these accounts. I will not delve into the specifics of the wide variety of published and unpublished accounts, but focus on their broad

assumptions and their implications for the phonetics of Dutch regressive voicing assimilation, and in particular the behaviour of three-term obstruent clusters. A detailed dissection of recent generative approaches to laryngeal neutralisation and voicing assimilation that is directly applicable to many recent analyses of Dutch is provided in the next chapter, and the reader interested in yet more detail is referred to sources such as [Trommelen & Zonneveld \(1979\)](#), [Booij \(1981\)](#), [Berendsen \(1983\)](#), [Booij \(1995\)](#), and [Ernestus \(2000\)](#). The second part of this section describes what the phonetic manifestations of Dutch RVA should be if it is to be regarded as a purely coarticulatory process, whilst [7.1.3](#) reviews earlier observations concerning the behaviour of three-term clusters.

7.1.1 Predictions of the ‘standard’ analysis

The three rules in [\(23\)](#) are the essential components of what I will call the ‘standard’ analysis of Dutch final neutralisation and regressive voicing assimilation. [\(23a\)](#) expresses the idea that final neutralisation is an asymmetric rule that changes lax obstruents into tense ones word finally. [\(23b\)](#) represents the idea that Dutch regressive voicing assimilation is both manner asymmetric (it is triggered only by plosives) and tense asymmetric (it is triggered only by lax obstruents). Whilst most accounts assume that the manner asymmetry is an idiosyncrasy of Dutch, there has been a tendency in recent generative work to view RVA as typically [tense]-asymmetric (according to this approach languages such as Hungarian and Yiddish belong to a marked or exceptional type: cf. [chapter 8](#)). Rule [\(23c\)](#) finally, expresses the idea that despite its similarities to fricative devoicing processes found elsewhere, Dutch postobstruent fricative devoicing reflects true linguistic process that spreads [+tense] rightwards rather than a passive devoicing process (cf. [4.3](#)).

The standard analysis of Dutch final neutralisation and voicing assimilation derives a number of predictions about the phonetic manifestation of these processes. In addition, depending on the relative ordering (or ranking) of the rules in [\(23\)](#) and the precise definition of the RVA rule, it generates a prediction about the (non)-application of RVA in three-term clusters with a medial fricative. First, as any lexical feature analysis, it predicts that Dutch regressive voicing assimilation should apply to all phonetic features involved in the realisation of [\pm tense] (cf. [prediction 10b](#) in [section 4.1.1](#)). Second, because it subscribes to a fortition account of final neutralisation, the standard analysis predicts that Dutch RVA is [tense]-asymmetric phonetically as well as phonologically. This means that Dutch word-final obstruents should be phonetically identical before fortis obstruents and sonorants: they are [+tense] in both environments and not subject to regressive assimilation (except in trivial fashion).

(23) The standard analysis of Dutch final neutralisation, RVA, and postobstruent fricative devoicing

a. Final neutralisation as fortition (see chapter 3)

$$\begin{bmatrix} -\text{son} \\ -\text{tense} \end{bmatrix} \rightarrow [+tense]_{-}\#$$

b. Dutch regressive voicing assimilation (e.g., Cohen et al. 1972; Trommelen & Zonneveld 1979; Berendsen 1983)

$$\begin{bmatrix} -\text{son} \\ +\text{tense} \end{bmatrix} \rightarrow [-tense]_{-}\begin{bmatrix} -\text{son} \\ -\text{cont} \\ -\text{tense} \end{bmatrix}$$

c. Postobstruent fricative devoicing (e.g., Trommelen & Zonneveld 1979; Lombardi 1999)

$$\begin{bmatrix} -\text{son} \\ +\text{cont} \\ -\text{tense} \end{bmatrix} \rightarrow [+tense]_{-}[-\text{son}]_{-}$$

This second prediction of the standard analysis is rarely questioned in the literature on Dutch regressive voicing assimilation and often seems to be regarded as part of the description of the process instead of as part of its analysis. Worse, it has become something of a self-fulfilling prophecy since it is generally left untested by instrumental studies of Dutch. The logic of this state of affairs is made explicit by the following passage from Slis (1986) (emphasis mine):

Since for syllable-final obstruents a final devoicing rule holds (Trommelen & Zonneveld, 1979; Booij, 1981), the first consonants of the clusters at issue (C_1) will have to be voiceless. This restriction implies that the second consonant in our clusters (C_2) has to be a voiced [i.e., lenis] obstruent; *if it was voiceless the clusters would consist of two voiceless obstruents in which no assimilation of voice could be studied.* (Slis 1986:313)

The predictions of the standard analysis concerning the behaviour of three-term clusters with a medial fricative are less unequivocal. On purely logical grounds, there are 4 possible surface (phonological) forms for the /ts/ sequence in a form such as /fi:ts/ + /band/, *bicycle tyre*: [ts], [tz], [ds], and [dz]. Depending on assumptions about rule ordering (or constraint ranking) and the precise formal definition of the regressive assimilation rule, the standard analysis can be set up to derive all of these sequences except [ds]

Consider first a procedural model that orders (23b) before postobstruent fricative devoicing. This type of model predicts that such forms are realised with a voiceless ([+tense]) medial fricative ([fi:tsbant]) because the former rule feeds the latter. The same is true if the rules are interpreted as violable well-formedness conditions or filters and (23b) dominates (23c). If the order of rule application (or ranking) is reversed it is predicted that /fi:ts/ + /band/ is realised as [fi:tzbant], because in this case (23c) feeds regressive voicing assimilation.

Note that given the definitions in (23), it is impossible to derive a form with a fully voiced three term cluster ([fi:dzbant]) since even if the medial /s/ in the cluster surfaces as [z], further rightward spreading of voicing is prohibited by the manner restriction on (23b). By contrast, obstruent + stop + lenis stop sequences, as in the phrase /kɔləkt bɛlən/ (*to make a collect call* or /ɑxt/ + /bɑ:n/, *rollercoaster*, are predicted to surface with *iterative* RVA, i.e., with voicing throughout, under either ordering or ranking of (23b) and postobstruent fricative devoicing.

A different situation emerges if (23b) is reformulated as an iterative rule that spreads [-tense] leftwards from both plosives and fricatives. Procedural models incorporating such a manner-symmetric regressive assimilation rule predict that all clusters ending in a lenis plosive end up fully voiced (cf. Booij 1981). In such models (23c) has to be ordered before RVA in order to rule out assimilation to lenis fricatives in word-initial position. The derivation of the surface form for /fi:ts/ + /band/ then proceeds as in (24): first, fricative devoicing fails to apply (or applies vacuously) to the cluster-medial /s/ in /fi:ts/ + /band/. This fricative is subsequently voiced by RVA, and because there are no longer any manner restrictions on the process, it can ‘transmit’ voicing to the preceding coronal stop by means of a second iteration of the assimilation rule.¹

- (24) Derivation of surface voicing in obstruent + fricative + lenis plosive clusters using a manner-symmetric RVA rule

Underlying form	/fi:ts/ + /band/
Final neutralisation	/fi:tsbant/
Fricative devoicing	N/A
RVA, iteration 1	/fi:tzbant/
RVA, iteration 2	/fi:dzbant/
Surface form	[fi:dzbant]

A third surface form for the final cluster of /fi:/ in /fi:ts/ + /band/ is pre-

¹Final neutralisation is incorporated for the sake of completeness. The OT analysis of Dutch laryngeal phonology in Grijzenhout & Krämer (1998) also predicts that the medial cluster of /fi:ts/ + /band/, albeit on different grounds. The key to this prediction is that word-initial underlyingly [-tense] fricatives are treated differently from those that acquire this specification through assimilation.

dicted by at least one OT model of Dutch laryngeal phonology. Lombardi (1997) presents a constraint-based analysis of Dutch voicing assimilation built around the interaction of two constraints, *IdentOnset(Laryngeal)* and *AGREE*, which demands manner-symmetric regressive voicing assimilation (cf. section 8.2.5). To avoid assimilation to word-initial lenis fricatives both constraints are dominated by *FricVoice*, which stipulates that postobstruent fricatives be [+tense] (hence voiceless). The result of this ranking is that all-tense sequences emerge as the optimal candidates for underlying obstruent + fricative + lenis stop clusters, because *FricVoice* filters out all candidates with a [-tense] medial fricative and *AGREE* banishes all remaining output forms with mixed voicing. In other words, it predicts that /firts/ + /band/ surfaces as [firtspand].

7.1.2 Predictions of the phonetic theory

The two most important predictions of a coarticulation-based theory of RVA with regard to the behaviour of Dutch word-final stop + fricative clusters are set out in (25a) and (25a) (the prediction about the phonetic manifestation of RVA is as before, cf. 25c). First, such clusters are predicted to be subject to regressive assimilation when followed by an actively (de)voiced obstruent. The effect of assimilation on a sequence of obstruents may not be fully proportional to that on a single stop or fricative if anticipatory articulation of the gestures involved in active (de)voicing starts relatively late (and is therefore relatively weak around the time of the onset of the first obstruent in the sequence). Furthermore, if one or both of the obstruents in a word-final clusters are actively devoiced, the combined coarticulatory ‘weight’ of their voicing targets may temper the effect of a following actively voiced lax obstruent. But coarticulation can not be switched off in the way a phonological assimilation rule can fail to apply, and as a consequence, the phonetic theory of RVA predicts that Dutch word-final obstruent clusters should be subject to some degree of assimilation.

- (25) Predictions of a coarticulation-based theory of voicing assimilation concerning the behaviour of word-final plosive + fricative clusters
- a. Dutch word-final plosive + fricative clusters are subject to regressive voicing assimilation
 - b. If these word-final clusters are phonetically underspecified for [tense], regressive assimilation is [tense]-symmetric: a following lax stop should trigger an increase in voicing relative to a baseline sonorant environment whilst a following tense obstruent should cause a decrease (cf. Ernestus 2000)
 - c. The assimilatory effects of fortis and lenis plosives are limited to voicing and features mechanically dependent on the production of voicing distinctions. Cf. (11a)

According to the model proposed by Ernestus (2000), Dutch neutralised obstruents are phonetically underspecified for [tense], which entails that they lack phonetic targets for voicing, and regressive voicing assimilation is a purely phonetic phenomenon. Ernestus notes that one unexplored prediction of this model is that contrary to the standard view, Dutch regressive assimilation at word boundaries should be (observably) [tense]-symmetric. I briefly sketched the reasoning behind this prediction in 4.1.2 in the context of the assimilatory behaviour of sonorants. The key assumption is that *ceteris paribus*, [0tense] obstruents have longer voiced intervals when preceded by a vowel than the corresponding [-tense] obstruents. In the former voicing continues from the vowel into the constriction phase of an obstruent until the transglottal pressure difference falls below the critical 200 Pa threshold or the glottis opens for some reason that is unrelated to voicing control. The latter are assumed to be accompanied by active devoicing gestures (glottal abduction, glottal constriction) that force vocal fold vibration to terminate at some time prior to the point of passive devoicing. Now if a [0tense] obstruent is followed by a sonorant, the duration of its voiced interval will be in accordance with the window for passive voicing as dictated by vocal tract aerodynamics and mechanics, because Dutch sonorants lack voicing targets and are therefore unable to exert any coarticulatory pressure on a preceding obstruent. The active devoicing gestures accompanying a [-tense] obstruent on the other hand, should be coarticulated during a preceding [0tense] fricative or stop. As a result, the [0tense] obstruent will itself become actively devoiced to some degree, and at least in principle, this will lead to voicing offset prior to the point determined by passive voicing alone. The same logic predicts that if a [0tense] obstruent precedes an actively voiced [-tense] obstruent, the length of its voiced interval is predicted to increase beyond the length of the passive voicing window.

Whether the predicted three-way pattern of voicing in neutralised obstruents before fortis obstruents (short voicing interval), sonorants (intermediate) and actively voiced lenis obstruents (long) indeed materialises hinges on the actual length of the passive voicing window for particular obstruents with a particular preceding context. Recall from 2.1 that estimates for the amount of passive voicing in a postvocalic obstruent run between 25 and 100 ms. The lower bound of this range is approximately the same as that of the presumably actively devoiced fortis velar stops of English and Hungarian in baseline and _[+tense] environments. This would suggest that if 25 ms was the approximate passive voicing window for (Dutch) neutralised obstruents, the coarticulatory effect of a following [-tense] stop would be impossible to observe in the speech signal (although it would be no less real at the articulatory level). However, if the passive voicing window for a neutralised obstruent is, say, 35 ms or longer, the three-way voicing pattern predicted by the phonetic theory of RVA should emerge from the

acoustics.

7.1.3 Observations on assimilation in three-term obstruent clusters

The assimilatory behaviour of clusters composed of more than two obstruents has never played a great role in the modelling of RVA, and that is perhaps part of the reason that observations on this point are relatively rare. Consequently, it is hard to determine whether RVA is typically iterative or not, and whether the manner of articulations of the obstruents involved typically imposes restrictions on iterativity. Nevertheless, Katz (1987) describes Yiddish regressive voicing assimilation as applying to all obstruents preceding the trigger. Thus, he transcribes the realisation of /ɛrft/ + /gəʃen/, *just happened*, as [ɛrʒdgəʃen]. Siptár & Törkenczy (2000) claim that Hungarian voicing assimilation is iterative, too. /list/ + /bø:l/ *from flour* for example, is said to surface with a fully voiced medial cluster: [lizdbø:l]. Opinions seem to differ with regard to Frisian: Riemersma (1979) appears to claim that no assimilation occurs in obstruent + fricative + lenis stop sequences, whereas examples provided by van der Meer & de Graaf (1986) indicate that such clusters are subject to RVA like any other sequence of obstruents in Frisian. The data from Yiddish, Hungarian, and Frisian therefore largely supports a phonetic view of regressive voicing assimilation, or, in light of the results presented in the previous chapter, a view of RVA as ultimately grounded in a phonetic process.

However, regressive voicing assimilation in Dutch has traditionally been described as non-iterative in obstruent + /s/ + lenis stop sequences. For example, Brink (1975), citing earlier work on Dutch, states that /firts/ + /bænd/ can be pronounced with a prevoiced lenis stop as in [firtsbænd], or with a (partially) devoiced lenis stop as in [firtsb̥ænd/], but is never realised with any voicing in the obstruents preceding the lenis stop. The status of this description (in part of the research community) as incontrovertible fact is reflected by Camminga & van Reenen (1980) who criticise the model developed by Booij (1981) for predicting RVA in stop + fricative + lenis stop sequences. It ostensibly solves the little rule ordering puzzle sketched above, indicating that the rule in (23c) takes precedence over (23b). At the same time, this description of Dutch regressive voicing assimilation raises doubts about a phonetic account of regressive voicing assimilation.

7.2 Methods

Subjects Subjects were 4 native speakers (MJ1, GBP3, both male, ER2, LB4, female) of Dutch between 21 and 45 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. Although all speakers were residents of the town of Groningen, where the local dialect is of the aspirating rather than the voicing type, this did not apply to the subject's speech, which can be roughly described as standard with minor (northern and western) local features.

Materials The stimuli for this experiment consisted of clusters combining an initial /p/ C₁, and a medial /s/ C₂ followed by a/p, t, b, d, m, h,/ C₃ or an unreduced vowel (/V/), which is usually preceded by a [ʔ] in Dutch.² Although there is evidence that final laryngeal neutralisation is complete in Dutch (Baumann, 1995), C₁ obstruents were consistently /p, k/ and orthographic <p, k>, to avoid a potential bias due spelling pronunciations or other incomplete neutralisation effects (cf. Fourakis & Iverson 1984 and chapter 3 above).

(26) Sample stimuli for experiment 3

- a. Het was Jaaps tunnel die onder water stond, niet zijn kelder
/hət vɑs ja:ps tʏnəl di: ɔndər vɑ:tər stɔnd ni:t zən keldər/
It was Jaap's tunnel that under-water stood, not his basement
It is Jaap's tunnel that was flooded, not his basement
- b. Het was een Kaaps meisje dat de hoofdprijs won, niet een Kaaps jongetje
/hət vɑs ən ka:ps mɛisjə dɑt də hɔ:vdpri:s vɔnni:t ən ka:ps jɔŋətjə/
It was a Cape-ADJ. girl-DIM. who the head-prize won, not a Cape-ADJ. boy-DIM
It was a little girl from the Cape who won the first prize, not a little boy from the Cape

/m/ functioned as baseline environment in the usual way. /h/ and /V/ were included as C₃ environments to establish the effects of independent glottal articulations on the voicing of a preceding obstruent cluster. Neither /h/ nor /V/ or the glottal stop normally preceding it can be characterised as [±tense] but the articulatory gestures involved in the production of [h] (glottal abduction) and [ʔ] (glottal constriction) are commonly found as active devoicing gestures in the production of plosives. Consequently, the phonetic theory of RVA predicts that the amount of voicing during word-final /ps/ clusters should be highly similar before /h/, /V/, and /p, t/ (Ernestus, 2000).

C₁s were embedded in proper name N₁s consisting of a single syllable and preceded by a long low unrounded vowel /a:/. The medial /s/ always represented an adjectival marker, as in /ka:p/ + /s/, *of, from, pertaining to the Cape*

²There was an additional set of stimuli combining an initial /k/ and medial /s/ with a /p, t, b, d/ C₃. The responses to these stimuli are excluded from the discussion below because they follow exactly the same pattern as the responses to the /ps/ + C₃ sequences.

or a (related) possessive marker as in /ja:p/ + /s/, *belonging to Jaap*.³ The carrier words (N₂) for C₃ were disyllabic nouns with an initial lexical stress. C₃ always preceded a long vowel or (phonotactically long) diphthong. The carrier words (N₂) for C₃ were disyllabic nouns with an initial lexical stress. The N₁ + N₂ collocations were further embedded in carrier sentences designed to attract a contrastive nuclear accent on N₂. Some sample stimuli (orthographic and phonological representations) appear in 26. Target clusters are represented in a slanted font.

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 read the list of stimulus sentences 3 times. For the first, *Normal* reading, the subjects were asked to read the stimulus items at a self-selected comfortable rate. In an attempt to simulate a noisy environment, the subjects were then fitted with sound-treated headphones conveying a 80 dB white noise signal (a noise level roughly comparable to that on a moving city bus) for the second reading, and asked to speak in such a way that they could understand their own speech. The aim of impoverishing the subjects' auditory feedback was to elicit a more hyperarticulated speech variety that is sometimes referred to as the *Lombard reflex* (Lombardi 1991; see Junqua 1996 for an overview). Henceforth the second reading task will therefore be referred to as the *Lombard* condition. For the third, *Fast*, reading, subjects were asked to read the stimulus items as fast as possible in order to create a bias to more hypoarticulated speech.

The three reading tasks or conditions were intended to elicit the same stimulus items on a 3-point hypoarticulation scale, so as to build up a relatively complete 'phonetic map' of the realisation of Dutch three term clusters and to increase the chances of observing any form of regressive voicing assimilation.⁴

With the exception of the use of impoverished auditory feedback to elicit 'clear' speech, this methodology is similar to methods used in a number of experimental studies of speaking rate effects on plosive VOT. Sometimes test subject are simply asked to produce the same set of stimulus items in fast or 'clearly enunciated' speech (e.g., Kessinger & Blumstein 1997), other experiments (Miller et al., 1986; Magloire & Green, 1999) use so-called magnitude production techniques, which essentially consist of instructions to test subjects to speak *n* times faster and slower relative to some self-selected 'normal' baseline.

During each of the three readings subjects were asked to repeat an item if

³Strictly speaking it is not clear whether these morphemes are [+tense] /s/ or [0tense] /S/, but nothing below crucially hinges on this. For typographical reasons I will represent them as /s/.

⁴Predictions about the behaviour of RVA in different global speaking registers are reviewed in the introduction to 7.2.2.

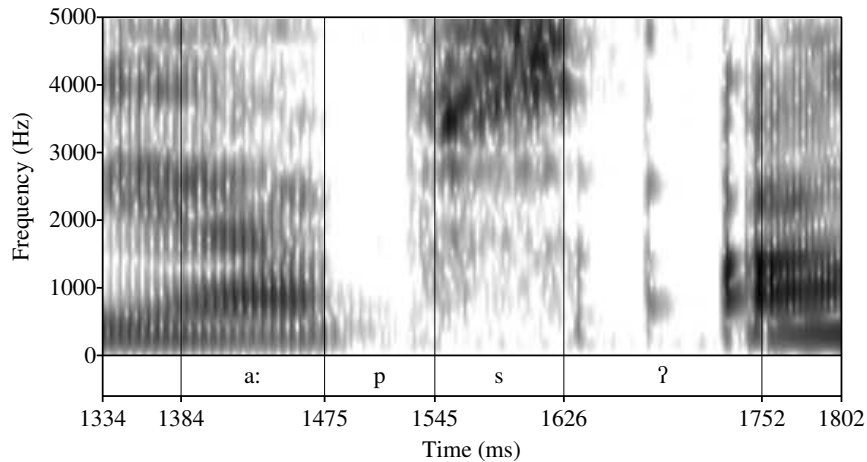


Figure 7.1: Experiment 3: sample segmentation of glottal stop preceded by /ps/. Broad band spectrogram of a /ps/ + [ʔ] cluster: subject = ER2, condition = Normal.

they produced a hesitation or speech error that was clearly audible to the experimenter and that affected the target cluster. In total, 1 ($C_1 = /p/$, $C_2 = /s/$) * 7 (C_3) * 10 (stimuli) * 3 (conditions) * 4 (speakers) = 840 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. 31 utterances had to be discarded because they contained a pause between C_2 and C_3 or small speech errors, leaving 809 utterances for segmentation and analysis.

The segmentation protocol was as for experiments 1 and 2 with added provisions for /m, h, V/. The boundary between /m/ a following vowel were determined on the basis of the offset of the nasal formant, whilst the offset of /h/ was defined as the onset of high frequency energy carried by the periodic source. [ʔ] was only marked as such if there was evidence of irregular glottal pulsing in the signal (this was virtually always the case): for an example, see figure 7.1. For reasons of time, V_1 was initially segmented and measured only for clusters beginning in /ps/ and ending in /p, t, b, d/. The results of the measurements was such that it was deemed unnecessary to investigate V_1 duration for clusters beginning in /ks/, or those ending in /m, h, V/.

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 7.1, ordered by speech segment. In the light of the previous chapter the ra-

Table 7.1: Acoustic measurements and derived measures for Experiment 3.

V ₁		Segment			V ₂		
(a)	Duration	(d)	C ₁ C ₂ Duration	(g)	C ₂ Closure dura- tion (stops)	(i)	F ₀ 10-50 ms after C ₁ offset
(b)	F ₀ 50-10 ms before C ₁ on- set	(e)	Voicing dura- tion	(h)	VOT (stops)		
(c)	F ₁ 50-10 ms before C ₁ on- set	(f)	Voicing ratio				

tionale for the measures summarised in table 7.1 should require little further comment. Note that all measurements in the column for C₁C₂ were performed twice, for /p/ and /s/ individually. Because there was no variation in C₂ context, no attempt was made to segment the release of /p/ from the preceding closure phase. As stated above, V₁ duration was only measured for clusters ending in a plosive C₃.

7.2.1 Main results

Phonetic features of C₃ plosives As can be gleaned from table 7.2 and figure 7.2, the contrast between the tense and lax C₃ plosives is as would be expected from a voicing language. /p, t/ have a small positive VOT that is comparable to the value for English /d/ found in experiment 1. The average VOT for /b, d/ is -54 ms, which is somewhat larger than the value reported for Hungarian reported in the previous chapter, but note that the average duration of oral closure is considerably longer too for the Dutch stops, so that the mean proportion of oral closure that is voiced (.55) is lower than the value found for Hungarian (.74). The difference in F₀ between tense and lax stops is comparatively large for both the female and male subjects: at 10 ms into the vowel the gap is 44 Hz (274 vs. 230 Hz) for the former and 31 Hz (214 vs. 183 Hz) for the latter. As in English and Hungarian, the sonorant baseline environment patterns with the lax rather than with the tense stops.⁵

These differences in VOT and F₀ stand up to statistical scrutiny. A t-test on the VOT data reveals a highly significant effect, $t(465) = 23.24$, $p < .001$ and the same applies to a three-way ANOVA for C₃ laryngeal specification on the F₀ values at 10 ms into the following vowel (female speakers only, clusters ending in /h, V/ excluded): $F(1,298) = 31.55$, $p < .001$. Tukey and Scheffe post-hoc tests show that as in English and Hungarian, the tense stops are distinct from

⁵Clusters ending in /h, V/ are excluded from figure 7.2 for clarity, and because segmentation of the offset of [h, ?] was felt to be relatively unreliable.

Table 7.2: Experiment 3: closure duration and VOT for C₃ plosives. All values in ms, and pooled across places of articulation (labial and alveolar) and reading tasks. Standard deviations in brackets.

C ₃	VOT	Closure duration	N
Tense (/p, t/)	16 (9)	89 (39)	234
Lax (/b, d/)	-54 (45)	85 (33)	233

both the lax stops and sonorant /m/ ($p < .001$ for all pairwise comparisons), whilst the means for the latter two groups are not significantly different.

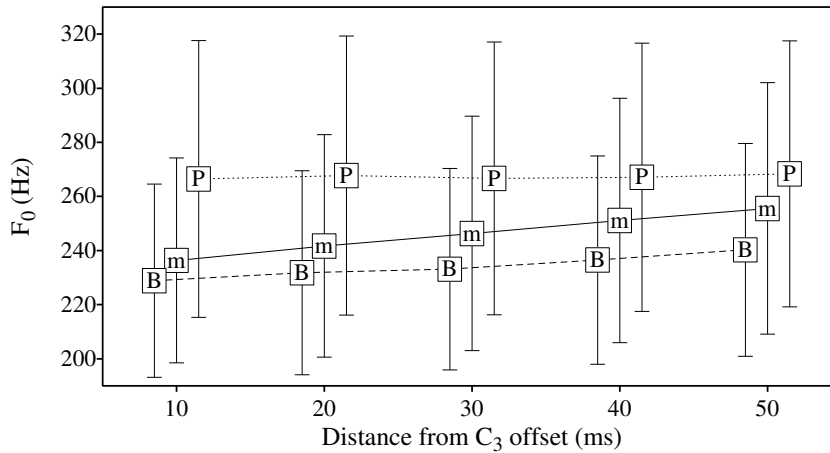


Figure 7.2: Experiment 3: F₀ (Hz) 10-50 ms into the vowel following C₃, for female speakers only. Clusters ending in /h/ and /V/ excluded, and data pooled over reading tasks. Error bars represent the mean \pm 1 standard deviation.

C₁ + C₂ voicing and duration Voicing and duration data for /ps/ clusters, which are summarised in figure 7.3, lend virtually unequivocal support to an articulation-based view of RVA on all counts. First, there is a clear increase in the duration of voicing before lax stops vis-à-vis all other contexts, which means that word-final /p/ + /s/ clusters are subject to regressive voicing assimilation under any viable phonetic definition of the term. Second, the duration of the voiced interval before /m/ is roughly intermediate between the durations of voicing before tense and lax stops, which indicates that, contrary to the received view, Dutch RVA is tense-symmetric. Third, /h/ and (preglottalised) /V/ pattern with tense stops as far as the voicing of a preceding obstruent cluster is

concerned, which supports the idea that RVA is largely a matter of the coarticulation of gestures involved in the realisation of voicing targets. Fourth, voicing and segmental duration do not maintain the inverse pattern that is typical of the phonetic expression of [tense] outside assimilation contexts. Variations in segmental duration rather seem to reflect mechanical interactions between glottal articulations involved in the expression of laryngeal (segmental) contrast and those involved in the production of fricatives.

Recall that according to the standard view of Dutch regressive voicing assimilation, there should be no difference in voicing between obstruents preceding a sonorant consonant and those followed by a tense obstruent. The (overall) voicing durations depicted in 7.3 plainly contradict this view and support prediction (25b) of the phonetic theory because there is a 13 ms difference between the two contexts. Note that the increase of voicing before lax stops is itself at odds with the assertion by Brink (1975) and others that no assimilation takes place in three-term clusters with a medial fricative. Interestingly, the effect of a fortis stop on the voicing of a preceding /ps/ cluster is highly similar to that of a preceding /h/ or preglottalised vowel. Given that glottal abduction and glottal compression are known active devoicing strategies this observation is consistent with the idea that the tense stops of voicing languages (in contrast to the lax stops of aspirating languages) are actively devoiced.

At first sight, the overall duration of /ps/ before lenis plosives, fortis plosives, /m/ and /h/ looks consistent with a representation of RVA as phonological feature spreading. $C_1 + C_2$ segmental duration is relatively long before /p, t/ (142 ms), short before /b, d/ (120 ms), whilst /m, h/ represent a more or less intermediate class. However, this classification of C_3 does not match the grouping implied by the voicing data, which classifies /h/ with the fortis stops rather than with /m/. Moreover, the standard theory places /V/ in the set of 'neutral' contexts (along with tense stops, /m/, and /h/), but it shortens the duration of a preceding /ps/ sequence even more than the lenis plosives. In other words, the inverse patterning of segmental duration and voicing that is predicted by a lexical feature analysis of RVA does not hold for the data reported here. This 'mismatch' between C_1C_2 voicing duration and segmental duration, is reminiscent of the behaviour of English and Hungarian C_1 as reported above.

This argument is bolstered if the segmental durations of /p/ and /s/ are considered separately. The right panel of figure 7.3 shows that differences in overall C_1C_2 segmental length are mainly due to differences in the duration of /s/ rather than the initial /p/: the means for the latter all cluster within a 8 ms band, whereas the maximal difference for the former is 20 ms. This pattern is familiar from the behaviour of C_1 plosive release duration and C_1 fricative duration observed in the results of experiment 2 (cf. table 7.3) and suggests an explanation along similar lines. In other words, it appears likely that mechanical

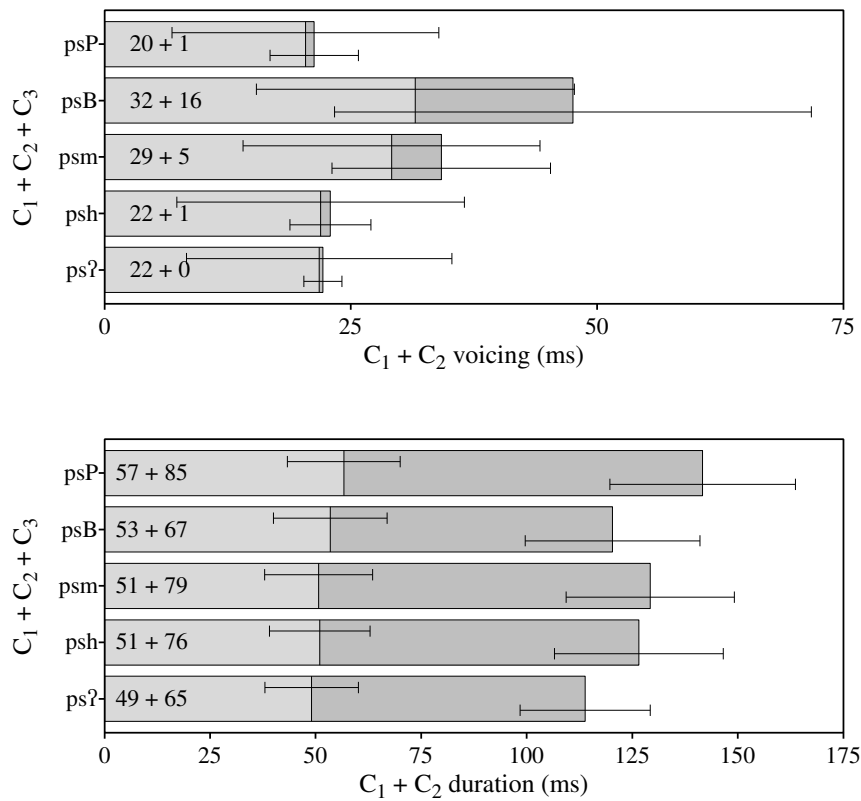


Figure 7.3: Experiment 3: voicing and duration of $C_1 + C_2$. Voicing (left) and segmental duration (right) of /p/ (bottom segments, dark grey fill) and /s/ (top segments, light grey) before fortis stops, lenis stops, /m/, /h/, and /N/ (all values in ms). Data pooled across registers.

interactions between the (glottal) articulations of C_3 and the medial /s/ form at least part of the explanation for the segmental duration facts.

At several previous points I have invoked the purely mechanical account of the linkage of duration and voicing in fricatives as proposed in [Stevens et al. \(1992\)](#): the vocal fold adduction required for the production of voicing inhibits the high transglottal airflow required for the production of frication noise, and consequently the frication phase of a fricative shortens if active voicing measures are imposed on it (e.g., by coarticulation). This idea extends to the shortening of a fricative before [ʔ] because the latter is also produced with an adduction gesture that inhibits airflow across the glottis. As noted above, it is the same gesture that impedes voicing (in a preceding sound) because it is stronger than

the adduction involved in modal voicing and leads to glottal compression. Thus, glottal coarticulation offers a natural account for the assimilatory behaviour of /s/ followed by [ʔ], which is rather Janus-faced from the perspective of a lexical feature analysis.

To test whether the above impressionistic observations stand up to statistical analysis, a number of tests were performed. First, clusters followed by a plosive C_3 were directly compared by means of a t-test on the voicing duration data, in order to demonstrate that RVA takes place even from the perspective of the standard view. This t-test indicates that the differences in C_1C_2 voicing duration is statistically highly significant, $t(465) = -10.73$, $p < .001$, and thus leaves little doubt about the inaccuracy of descriptions of Dutch three-term clusters with a medial fricative as impervious to RVA.

Table 7.3: Experiment 3: results of Tukey and Scheffe post hoc tests on the ANOVAS for $C_1 + C_2$ voicing (top) and segmental duration (bottom). *: significant difference ($p < .05$) according to both tests; \diamond : significant difference ($p < .05$) according to Tukey only; n.s.: difference not significant on either test.

	/p, t/	/b, d/	/m/	/h/	/V/
/p, t/		*	*	n.s.	n.s.
/b, d/	*		*	*	*
/m/	*	*		*	*
/h/	n.s.	*	*		n.s.
/V/	n.s.	*	*	n.s.	

	/p, t/	/b, d/	/m/	/h/	/V/
/p, t/		*	*	*	*
/b, d/	*		\diamond	n.s.	n.s.
/m/	*	\diamond		n.s.	*
/h/	*	n.s.	n.s.		*
/V/	*	n.s.	*	*	

Next, 3 one-way ANOVAs for C_3 laryngeal specification were performed on the C_1C_2 voicing and duration data with the baseline environment /m/ and /h, V/ included as three separate laryngeal specifications in addition to [\pm tense]. The ANOVAs on the voicing duration data, $F(4,804) = 48.92$, $p < .001$, and segmental duration data $F(4,804) = 25.77$, $p < .001$, both yield highly significant effects. Results of Tukey and Scheffe Post Hoc tests are summarised in table 7.3. The top panel shows the pairwise comparisons for C_1C_2 voicing duration, which clearly supports the impressionistic grouping of C_3 contexts for this parameter as /p,t/, /h/, /V/ vs. /b,d/ vs. /m/: for example, mean C_1C_2 voicing durations before the [+tense] plosives, /h/, and /V/ are all significantly different from

those before both /b, d/ and /m/ but there are no such differences within the first group. The fact that the mean C_1C_2 voicing duration before /m/ is significantly different from the mean voicing durations before /p, t/ as well as /b, d/ indicates that Dutch RVA is indeed [tense]-symmetric.

The Tukey and Scheffe results for C_1C_2 segmental duration do not support the classification indicated by the left panel, however. As suggested by the right panel of figure 7.3, the clearest split here appears to be between /p, t/ and all other contexts. But no clear grouping emerges within the set of remaining C_3 environments: for example, whereas /b,d/ is distinct from /V/ in terms of voicing duration the difference in C_1C_2 segmental duration is not statistically significant. Thus, the statistical tests confirm the mismatch between voicing duration and segmental duration identified above (as well as in the English and Hungarian data), thereby casting yet more doubt on a lexical feature analysis of RVA which predicts that such mismatches should not occur.

Classifying target clusters as (perceptibly) ‘assimilated’ vs. ‘unassimilated’

In an experimental production study of regressive voicing assimilation in Dutch two-way obstruent clusters, Slis (1986) employs a technique of quantifying RVA that is different from the methods used so far in this study. Slis classifies all obstruents preceding a lenis plosive C_2 as ‘unassimilated’ if they have a VTT < 50 ms and as ‘showing regressive assimilation’ if their VTT exceeds 50 ms. The cut-off point is based on the VTT of singleton intervocalic stops in Dutch as measured by Slis (1970), which indicates that there is a probability < .0025 that fortis stops have a VTT equal or greater to the mean of 25 ms + 2 standard deviations (10 ms) + 5 ms = 50 ms.⁶ The definition of regressive voicing assimilation used by Slis (1986) assigns strict acoustic criteria to a method that is used in transcription studies and consequently his methodology exposes the size of the effect of [\pm tense] on C_1 voicing duration relative to the inherent variance within a laryngeal category in an intuitively transparent way. The relative magnitude of [\pm tense] effects vis-à-vis the noise caused by within-category variation has been used as a rough indicator of perceptual salience: O’Shaughnessy (1981) suggests that effects smaller than or equal to a single standard deviation from a baseline mean should be treated as below the threshold of perception.

There are several ways in which this technique can be applied to the C_1C_2 voicing data from the present experiment. Given that assimilation of C_1C_2 voicing appears to be [tense]-symmetric a natural procedure is to define three classes of /ps/ sequences using the overall mean C_1C_2 voicing duration of 31 ms and its standard deviation of 26 ms: a ‘voiceless’ category or ‘band’ with relatively short intervals, a ‘neutral voicing’ class centered around the overall mean, and a ‘voiced’ category with relatively long voiced intervals.

⁶I.e., provided that the sample size is sufficiently large and VTT is normally distributed.

Regardless of the precise settings of the boundary values delimiting the three categories, the picture of Dutch regressive voicing assimilation that emerges from this method is not substantially different from the one drawn above. Figure 7.4 depicts the classification of /ps/ clusters if the neutral band is defined as the overall mean of $31 \text{ ms} \pm 1$ standard deviation of 26 ms. There are very few ‘voiced’ clusters preceding [+tense] plosives, /h/, or /V/, whereas 30 % of the cases before /b,d/ belong in this class, with the ___/m/ context almost exactly halfway in between (15%). The frequencies of ‘devoiced’ clusters hint at the same natural classes of assimilation environments, with similar frequencies before /p,t/ and /h/ and /V/ (17, 18, and 15% respectively), and lower figures for /m/ (8%) and /b,d/ (3%).

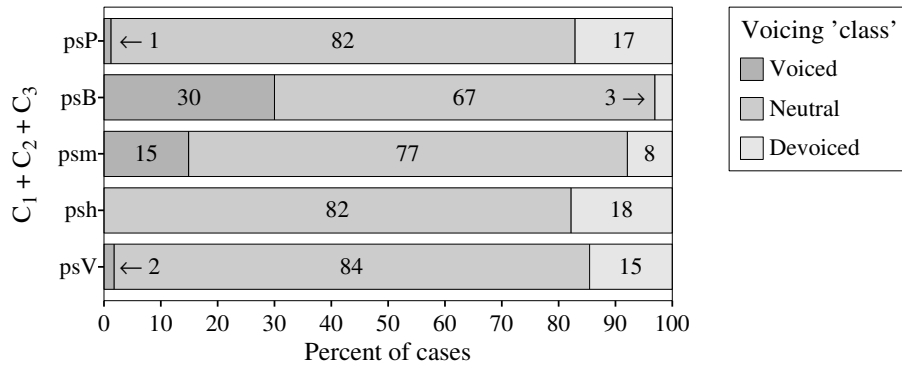


Figure 7.4: Experiment 3: frequencies of ‘devoiced’, ‘neutral’, and ‘voiced’ /ps/ clusters before fortis stops, lenis stops, /m, h, V/ if the neutral category is defined as 31 ± 26 ms (the overall mean ± 1 standard deviation). Numbers inside columns represent numbers of cases.

In his study of RVA in Dutch, Slis (1986) reports that lenis stops trigger 86% ‘regressive assimilation’ in preceding singleton plosives across a word boundary and before stress. This figure is considerably higher than the proportion of (equivalent) ‘voiced’ cases before [-tense] plosives in the present study. Note, however, that in the classification illustrated in figure 7.4 the cut-off point between the ‘neutral’ and ‘voiced’ bands is 57 ms of voicing as opposed to Slis’s 50 ms. If the cut-off point is lowered to 50 ms the proportion of ‘voiced’ cases before [-tense] plosives rises to 36%, which is identical to the frequency of ‘assimilated’ singleton fricatives found by Slis (in the relevant environment). This implies that, however real, the effect of RVA on plosive + fricative clusters are in some sense weaker than the effect on singleton plosives and therefore perhaps less audible. This may in turn account for the claims in the descriptive literature

that regressive voicing assimilation does not apply to plosive + fricative clusters in Dutch.

V₁ duration The mismatch between voicing duration and segmental duration in /ps/ sequences is consistent with prediction 25c of the phonetic theory. The same applies to absence of an assimilation reflex in V₁ duration. Vowel duration is known to pattern with [±tense] in the familiar way before word-medial obstruents in Dutch (e.g., Slis & Cohen 1969a), and can therefore not be dismissed as irrelevant by proponents of a lexical feature analysis. Figure 7.5 represents the mean duration of V₁ preceding /ps/ + [+tense] plosive and /ps/ + [-tense] plosive sequences (note again that V₁ was not segmented in any of the remaining cluster types for reasons of time). V₁ is slightly (3 ms) longer when C₃ is a lenis plosive, but a t-test, $t(463) = -0.82$, indicates that this difference is far from statistically significant. It appears therefore, that regressive assimilation in Dutch behaves much as its English counterpart in affecting obstruent voicing but not any of the other correlates of [tense].

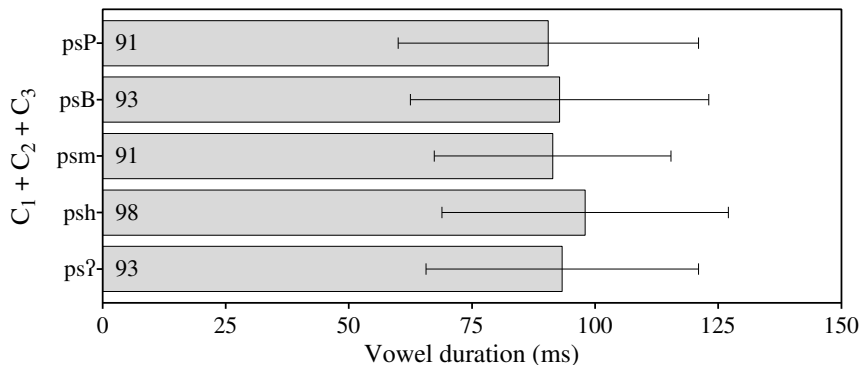


Figure 7.5: Experiment 3: mean duration (ms) of V₁ before /ps/ preceding tense and lax plosives. Error bars represent the mean + 1 standard deviation

Low-frequency spectral features F₀ and F₁ preceding /ps/ clusters mimic the behaviour observed for these features in the data from experiment 1. Despite the relatively large differences in F₀ following C₃ there is no regressive assimilation of fundamental frequency. F₁ values preceding C₁ on the other hand, do show some effect of C₃. Table 7.4 gives values for the first formant of the vowel /a:/ at 10 ms before the onset of C₁. Just as English /d/, Dutch /b, d/ appear to cause a decrease in the F₁ of the vowel preceding C₁, and the magnitude of the difference between tense and lax stops is similar in the two languages, too

(approximately 20 Hz). Likewise, the behaviour of Dutch /m/ mirrors that of English /r/: both sonorants pattern with tense rather than with lax obstruents. Note that the values for /h, V/ appear to be intermediate between the extremes defined by the lax stops on the one hand and sonorants and tense stops on the other. A one-way ANOVA for C_3 laryngeal specification on the F_1 values at 10 ms before the onset of C_1 (clusters ending in /h, V/ excluded) reveals an effect that is significant, $F(2,598) = 4.13$, $p < .02$, but considerably weaker than that obtained from experiment 1.

Table 7.4: Experiment 3: F_1 values (Hz) of the vowel /a:/ at 10 ms before the onset of C_1 . Standard deviations in brackets, and data pooled across reading tasks.

C_3	F_1 at C_1 - 10 ms	N
/p, t/	698 (94)	242
/b, d/	678 (87)	242
/m/	700 (77)	114
/h/	689 (81)	118
/V/	690 (95)	110

As noted in 5.4, the articulatory source of low-frequency spectral cues to [\pm tense] remains unclear, and consequently it is difficult to gauge the significance of assimilatory effects on F_1 .

7.2.2 Reading task effects

Descriptions often suggest that the degree of voicing assimilation between words somehow increases in ‘casual’ or fast speech or becomes applicable in a wider range of contexts, and there is a range of possible explanations for such observations, including a decrease in the number and length of physical pauses. However, it is difficult to state any precise predictions about the interaction of regressive voicing assimilation and global speech register without a highly formalised speech production model. Generative frameworks such as those of Kaisse (1985) and Nespor & Vogel (1986) tend to model rules that vary with speech register as ‘late’ and/or optional parts of a phonological derivation, or in terms of prosodic reanalysis. But such models are inherently committed to a phonological, all-or-nothing view of rule application and therefore shed no light on the question of how a (potentially) gradient assimilation process might be shaped by changes in global register.

In its present prose formulation, the model described in chapter 1 does not derive any precise predictions about the interaction between global register variation and regressive voicing assimilation either. Because hypoarticulation (con-

ceived as a relaxation of auditory targets) plays a key role in this model, there is a further complication in the interpretation of the present results if any variation induced by the three reading tasks cannot be characterised in terms of hypoarticulation. To the extent that the subjects' responses to the three reading tasks does show systematic variations in hypoarticulation, it might be expected that increased hypoarticulation would lead to three-term obstruent clusters converging on a completely passively voiced configuration (i.e., with an initial voicing tail and if the constrictions remain in place long enough, subsequent devoicing) regardless of their underlying specification. Note that this prediction runs counter to the common claim that assimilation increases in fast and 'casual' speech. However, a number of other factors including prosodic phrasing can be expected to distort this convergence. Moreover, even in the absence of such possibly confounding factors it might be difficult to identify any convergence on a passively voiced configuration, since global hypoarticulation is likely to affect segmental duration as well as the implementation of voicing contrasts.

In the following I will therefore simply describe the effect of the different reading tasks on the voicing of Dutch three-term obstruent clusters in the context of more global properties of the subjects' speech. Perhaps the most surprising part of this description is the observation that there is little evidence for an increase in regressive assimilation in fast speech. This finding contradicts earlier work on Dutch by [Menert \(1994\)](#).

Global effects of reading task Table 7.5 summarises the effects of reading task on a number of utterance-level variables: overall utterance duration, utterance mean F_1 , and utterance F_0 range. Utterance duration was determined on the basis of segmentation by hand, and can be interpreted as an indicator of overall speaking rate. Utterance mean F_1 was defined as the raw mean of the F_1 values of all voiced samples of an utterance. It was extracted automatically from the formant tracks produced by the Burg algorithm embedded in PRAAT 4.0. F_1 has been observed to increase in 'loud' and shouted speech, presumably because of a greater average jaw opening (hence a lower tongue height) during vowels under such conditions ([Rostolland, 1982](#); [Bladon, 1986](#); [Junqua, 1996](#)). Consequently, utterance mean F_1 was expected to be highest in the Lombard and lowest in the Fast readings. F_0 range was defined as the distance in Hz between utterance F_0 maxima and minima which were again extracted automatically, from the pitch tracks produced by the autocorrelation routine of PRAAT 4.0. F_0 range may be seen as a rough indicator of (perceptual) pitch range. Note that these are fairly noisy measures and the main motivation to use them was that they involved a minimum of additional hand labelling of speech samples.

All three measures could be said to exhibit register variation because they show different values for the 3 reading tasks, in a fashion that is consistent with

the literature on the topic. Broadly speaking, utterance duration, mean F_1 , and F_0 range all decrease between the Lombard and Fast conditions. The one exception is represented by the F_0 range for male speakers, which is 5 ms greater in the Normal than in the Lombard condition. However, perhaps the most conspicuous fact about the data summarised in table 7.5 is that the differences between the Lombard and Normal conditions are (considerably) smaller than those between the Normal and Fast conditions: there is a marginal difference (8 ms) in utterance duration between the Lombard and Normal conditions, as opposed to a 65 difference between the Normal and Fast conditions, a 20 (vs. 83) Hz difference in mean F_1 , and a 14 Hz (vs. 24) difference in F_0 range for the female speakers.

A series of ANOVAs lends support to the impression that the speech produced in response to the different reading tasks has different global characteristics, with the strongest contrast emerging between the Fast condition on the one hand and the Lombard and Normal readings on the other. A one way ANOVA for *reading task* (Lombard vs. Normal vs. Fast) on the utterance duration data shows that the effect is highly significant, $F(1,1264) = 422.66$, $p < .001$, but Tukey and Scheffe post hoc tests indicate that whereas the mean utterance duration for the Fast condition is significantly different from both the Normal and Lombard conditions (all $p < .001$), the difference between the means for the Lombard and Normal conditions is not significant. A one way ANOVA for *reading task* on the utterance mean F_1 data also yields a highly significant effect, $F(1,1264) = 358.18$, $p < .001$, but in this case Tukey and Scheffe post hoc tests show that all pairwise comparisons yield highly significant results, despite the relatively small difference between the Lombard and Normal conditions.

Table 7.5: Global phonetic effects of reading task: utterance duration (ms), utterance mean F_1 , and F_0 range (Hz) for female and male speakers (all combinations of /ps/ and C_3 included. Standard deviations appear in brackets.

Variable	Lombard	N	Normal	N	Fast	N
Duration	3553 (711)	426	3545 (651)	420	2480 (457)	421
Mean F_1	799 (58)	426	779 (57)	420	696 (62)	421
F_0 range (female)	210 (75)	214	196 (72)	211	172 (66)	210
F_0 range (male)	130 (26)	212	135 (32)	209	114 (53)	211

A two way ANOVA for *reading task* * *gender* (female vs. male) yields highly significant effects of *reading task*, $F(1,1261) = 27.86$, $p < .001$, as well as of *gender* $F(1,1261) = 422.59$, $p < .001$, and *reading task* * *gender*, $F(2,1261) = 4.72$, $p < .01$. The main effect of *reading task* evinces register variation even, in the F_0 range of male and female speakers pooled together. However, Tukey and Scheffe post hoc tests for *reading task* show that as for utterance duration, the means for the Lombard and Fast readings are not significantly different, al-

though both are significantly different from the mean utterance duration in the Fast condition (both $p < .001$). It seems likely that the behaviour of the male speakers, which produced a smaller F_0 range in the Normal than in the Lombard condition contributed to the lack of statistically significant contrast between these two conditions. The same phenomenon (and the reverse patterning of F_0 range in the female subjects' speech) may well underlie the significant interaction between *reading task* and *gender*. Finally, the main effect of *gender* attests to the greater pitch range, across registers, for the female speakers.⁷

Reading task effects on VOT in C_3 plosives The effects of reading task on the VOT of C_3 plosives was examined because the phonetic literature offers some good ground for (cross-linguistic) comparison.

It is a recurring observation in experimental studies on the topic that the VOT of short lag ([p, t, k] and [b, d, g]) plosives is impervious or only slightly sensitive to variations in speaking rate, whereas the VOT of long lag and prevoiced stops is sensitive to such variations (Miller et al., 1986; Pind, 1995; Kessinger & Blumstein, 1997; Magloire & Green, 1999). As visible in figure 7.6 the same applies to the present experiment: the mean VOT of the fortis plosives varies within a 2 ms window whereas there is considerable register-based variation in the prevoicing of the lenis class. The only result that is somewhat surprising is that lenis plosive VOT in the Lombard condition is smaller than in the Normal condition. Note however, that Magloire & Green (1999) report a similar reverse register effect on the VOT of Spanish fortis stops, which some of their test subject was produced with a longer VOT at faster rates. More generally, it appears that the effects on VOT of (elicitation techniques aimed at) slow speech are far less pronounced than those of fast speech (when compared with a 'normal' condition: cf. Miller et al. 1986; Magloire & Green 1999). Thus the results of the present experiment are more or less consistent with earlier studies.⁸

A two way ANOVA for *reading task* * C_3 laryngeal specification ([+tense] vs. [-tense]) reveals highly significant effects of *reading task*, $F(2,461) = 13.98$, $p < .001$, C_3 laryngeal specification, $F(1,461) = 607.49$, $p < .001$, and *reading task* * C_3 laryngeal specification, $F(2,461) = 16.98$, $p < .001$. The effect of C_3 laryngeal specification can hardly be surprising given the distinctly different VOT values for fortis and lenis plosives, while the effect of *reading task* indi-

⁷Recalculating F_0 on the basis of F_0 minima and maxima expressed in (base 100 Hz) *Semitones* has little effect on this outcome: the mean difference in F_0 range between male and female speakers is 2.77, 1.84, and 2.18 ST under the Lombard, Normal, and Fast conditions respectively.

⁸VOT was calculated as 0 - C_3 closure voicing for lenis plosives with closure voicing > 0 ms, and the interval between the onset of the release burst and voicing onset in ('devoiced') lenis plosives with no closure voicing and fortis plosives (all of which were fully voiceless). Data for the /ks/ + plosive clusters exhibits a VOT patterning that is nearly identical to that illustrated in figure 7.6.

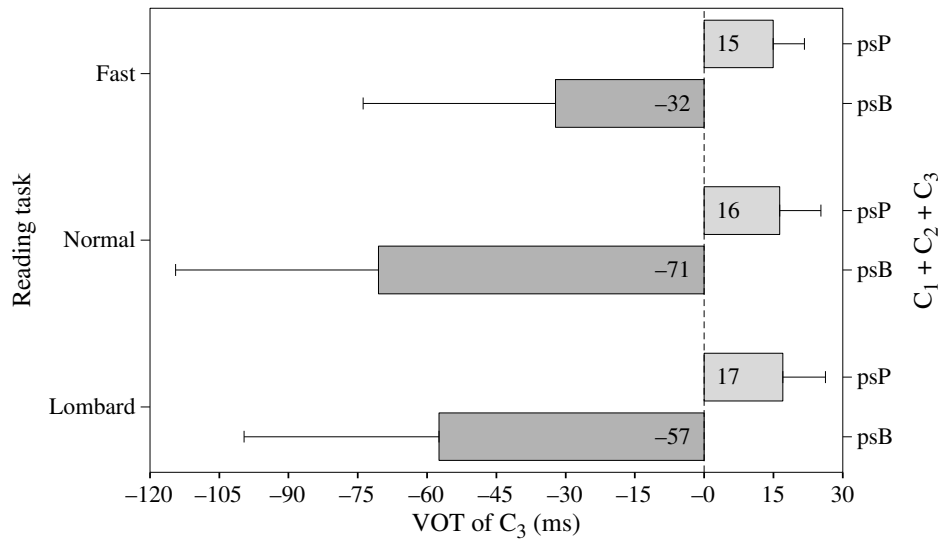


Figure 7.6: Experiment 3: effects of reading task on the VOT of [±tense] plosives preceded by /ps/. Error bars represent the mean \pm 1 standard deviation.

cates that there is a statistically significant amount of register variation in the realisation of VOT. However, the interaction between *reading task* and *C₃ laryngeal specification* implies that the main effect of *reading task* may be almost fully due to the variation of the lenis stops. Moreover, Tukey and Scheffe post hoc tests for *reading task* show that the mean Fast VOT is significantly different from those in the Lombard and Normal two conditions, but that there is no significant difference between the latter two. This reinforces the conclusion that the change from Lombard to Normal speech affects the subject's speech (and VOT in particular) far less than the change from Normal to Fast speech.

The interaction between register and RVA In the light of the foregoing it is somewhat surprising that the present experiment reveals little evidence for the idea that register or even plain rate have an observable effect on regressive voicing assimilation.

Figure 7.7 plots $C_1 + C_2$ voicing duration (as an index of RVA) against the combined duration of $C_1 + C_2 + C_3$ closure for clusters composed of /ps/ + /b, d, m, p, t/. sequences ending in /h, V/ were excluded for the sake of simplicity but nothing crucial hinges on this as they behave in a near-identical fashion to the clusters ending in a fortis plosive. $C_1 + C_2 + C_3$ closure duration was chosen as an index of hyparticulation (assuming that speech rate translates

into hypoarticulation) because it reflects the impact of the reading tasks at the same point in the signal where RVA occurs (but note that using any of the global measures discussed above leads to the same conclusions).

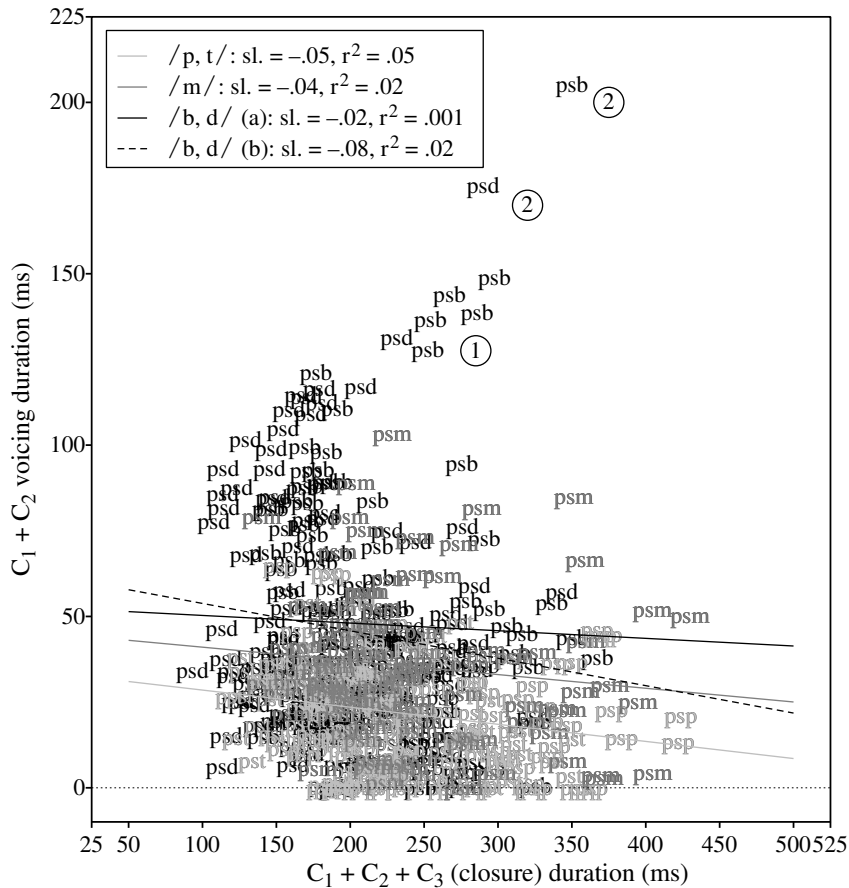


Figure 7.7: Experiment 3: scatter plot of $C_1 + C_2$ voicing duration (ms) against $C_1 + C_2 + C_3$ duration (ms) and results of linear regression analyses. The solid black line (a) represents the regression line for clusters ending in /b, d/ if all cases are included whilst the dashed black line represents the regression line for these clusters if the outliers labelled (2) are excluded.

Given that Dutch RVA is tense-symmetric, the most obvious measure of *degree of assimilation* is the difference in voicing between /ps/ clusters preceding lax stops and those preceding tense stops. An ‘increase’ in assimilation then means a divergence of C_1C_2 voicing values before tense and lax stops, and an ‘increase’ in assimilation as a result of higher speaking rate should appear as

a convergence of voicing values towards the left-hand side of the scatter plot in figure 7.7. In the introduction to this section I suggested a hypothesis which cast increased hypoarticulation as a 'leveller' of voicing distinctions in obstruent clusters and hence of distinctions transmitted backwards by means of regressive assimilation. This hypothesis predicts a *convergence* towards the left-hand side of the scatter plot (assuming again that speech rate may be used as an index of hypoarticulation).

However, no clear trends in either direction are visible in figure 7.7. A linear regression analysis indicates that across C_3 contexts, the amount of $C_1 + C_2$ voicing decreases with increasing $C_1 + C_2 + C_3$ closure duration (the slope of the regression line is $-.08$), and thus increases with degree of hypoarticulation, but although the effect is significant ($p < .001$), the reduction in variance, r_2 , is only $.03$. One plausible source for this increase in voicing in more hypoarticulated speech is increased coarticulation with the preceding vowel (see chapter 4). Separate regression analyses indicate a similar (significant) decrease in $C_1 + C_2$ voicing with increasing $C_1 + C_2 + C_3$ closure duration in clusters ending in [+tense] plosives, slope = $-.05$, $r_2 = .05$, $p < .005$, but fail to reveal effects for clusters with $C_3 = /m/$, slope = $-.04$, $r_2 = .02$, not significant, or with $C_3 = /b, d/$, slope = $-.02$, $r_2 = .001$, not significant. Thus, the slight convergence of the regression lines for these classes at lower values of $C_1 + C_2 + C_3$ closure duration, which is illustrated by the solid lines in figure 7.7, does not constitute evidence for the levelling hypothesis.

Moreover, the slope of the regression for clusters ending in $/b, d/$ is heavily influenced by the two outlier values labelled (2) in figure 7.7, and to a lesser extent by the small cloud of 6 tokens labelled (1). Impressionistic inspections of the recordings of these cases, which were produced by a single speaker (MJ1) reveal nothing remarkable to the ear of the author that would form independent grounds for excluding them from the database. Thus, there seems little ground for excluding the cases labelled (1): their voicing values are equal to or smaller than the $/ps/ +$ lenis plosive mean + 3 standard deviations (149), and are more or less contiguous with the upper $C_1 + C_2$ voicing ranges of the main cloud. But there does seem to be a case for treating the two tokens labelled (2) as genuine outliers: their $C_1 + C_2$ voicing durations represent the $/ps/ + /b, d/$ mean (47 ms) + 4.64 and 3.79 standard deviations (34 ms) respectively, and there are large gaps between the category (2) voicing values and the values for neighbouring tokens.

A regression analysis on the clusters ending in a lenis plosive with these two outliers excluded yields a weakly significant effect, $r_2 = .02$, $p = .05$ and a slope that is greater ($-.08$) than that for $/p, t/$, which in turn results in a small divergence between the regression lines for [+tense] and [-tense] in more hypoarticulated speech (cf. the dashed black line (b) in figure 7.7), as is predicted

Table 7.6: Experiment 3: effects of reading task on C₁C₂ voicing ratio (standard deviations in brackets).

C ₃	C ₁ C ₂ voicing ratio					
	Lombard	N	Normal	N	Fast	N
/p, t/	.12 (.09)	79	.12 (.10)	78	.25 (.15)	77
/b, d/	.36 (.27)	77	.29 (.23)	78	.61 (.33)	78
/m/	.23 (.15)	39	.24 (.20)	39	.39 (.20)	36
/h/	.16 (.10)	39	.17 (.11)	39	.23 (.16)	40
/V/	.19 (.09)	35	.16 (.13)	37	.25 (.15)	38

by optional rules models of regressive voicing assimilation. But note that the effect is a best a weak one: the proportion of variance that is accounted for by the regression lines for /ps/ + /p, t/ and /ps/ + /b, d/ does not exceed .08, even if the outliers labelled (1) are excluded from the analysis.⁹

In fact, the best evidence for an increase in regressive voicing assimilation comes from the marked increase in C₁C₂ voicing *ratio* before lenis plosives in the Fast condition, which is brought about by the concomitant increase in voicing and compression of segmental duration in the Fast condition. As table 7.6 illustrates, C₁C₂ voicing ratio increases across C₃ contexts in the Fast condition, but not to the same extent as before /b/ and /d/. A two-way ANOVA for *reading task* * C₃ *laryngeal specification* on the voicing ratio data (clusters ending in /h, V/ excluded) shows highly significant main effects of *reading task*, $F(2,572) = 47.65$, $p < .001$, and C₃ *laryngeal specification*, $F(2,572) = 87.47$, $p < .001$, as well as a significant interaction of *reading task* and C₃ *laryngeal specification*, $F(4,572) = 4.94$, $p < .005$. Whilst the main effects indicate that C₁C₂ voicing ratio is subject to effects of reading task and regressive assimilation, the interaction show that not all conditions contribute evenly to these effects. The sharp increase in voicing ratio before lax stops in the Fast condition is the most likely cause of this interaction. Provided that voicing ratio is perceptually relevant, an increased voicing ratio before /b, d/ may explain why RVA is perceived to increase with speaking rate.

7.3 Conclusions

With regard to the question of whether regressive voicing assimilation applies in Dutch obstruent + fricative + lax stop sequences, the findings of the experiment reported in this chapter seem to me fairly unequivocal. The voiced interval of

⁹Excluding the tokens labelled (1) and/or (2) does not affect the conclusions reached in the previous sections in any (statistically) significant way.

word-final /ps/ clusters is affected by the status of a following obstruent even if baseline environments are set aside, and this finding is at odds with traditional accounts of the behaviour of such sequences. From a typological perspective this result brings Dutch RVA in line with the processes found in Hungarian and Yiddish, which have long been described as iterative.

The observation that assimilation does occur in three-term clusters with a medial fricative is conform prediction (25a) but will in itself not rock major theories of laryngeal phonology. However, some of the other results reported above have more far-reaching ramifications. First and foremost, Dutch regressive voicing assimilation appears to be [tense]-symmetric rather than asymmetrically triggered by lax plosives, as held by the standard theory. This finding is consistent with a coarticulation-based view of RVA (25b), which predicts that both the tense and lax obstruents of voicing languages are able to trigger voicing assimilation. Perhaps more importantly, it is entirely consistent with the hypothesis in Ernestus (2000) that final obstruents in Dutch are phonetically underspecified for [tense]. Even if this hypothesis was not directly put to the test (this would involve comparing neutralised with non-neutralised obstruents in phonetically similar contexts) the observation that regressive assimilation to tense stops and [h, ʔ] is responsible for what is heard (by linguists) as final devoicing, suggests that the standard view of final neutralisation as fortition may be flawed.

Experiment 3 supplied further evidence for hypothesis (25c). V_1 duration nor C_1C_2 segmental duration nor F_0 microprosody pattern with C_1C_2 voicing, which is the only feature that displays clear traces of regressive voicing assimilation. This finding is entirely consistent with the results of experiment 1. As in Hungarian and English C_1C_2 segmental duration varies under the influence of the following context, but as in the case of Hungarian C_1 fricative duration and C_1 release burst duration it seems that the variations in question can to a large extent be explained in terms of glottal coarticulation between gestures involved in the production of voicing distinctions or [ʔ] and the glottal abduction required for the generation of turbulence noise in the oral tract. The data reported above also indicate that the somewhat puzzling [tense]-asymmetric effect of C_2 on the first formant of a vowel preceding C_1 is not limited to the British English of the subjects of experiment 1. The mechanism responsible for this effect remains unclear however, as does the answer to the question why F_0 microprosody should not also 'spread' backwards.

Finally, this chapter attempted to examine the effects of global register variation on the phonetic manifestation of RVA. It is commonly observed that voicing assimilation phenomena increase with increases in speaking rate, and Menert (1994) offers experimental evidence for this view. However, it is only supported by the results of the present experiment to the extent that there is an increase in the voicing ratio of /ps/ preceding a lax stop in the Fast reading condition. An

alternative hypothesis about the relation between speaking rate (as a bias towards a certain degree of hypoarticulation) and the realisation of voicing distinctions in obstruents holds that rate increases act to level underlying distinctions in voicing and hence to diminish the effects of RVA. However, no support was found for this hypothesis.

