PHONETIC CORRELATES OF POSTVOCALIC /r/ IN SPONTANEOUS DUTCH SPEECH

Leendert Plug

Abstract

This paper reports on a study of the realisation of postvocalic /r/ in stressed syllables followed by /t/ or /d/ in Dutch. Two recurrent observations regarding /r/ in this context are that /r/ can be 'deleted', and that /r/ has an effect on the quality and duration of a preceding vowel and a following alveolar consonant. Most recently, Plug & Ogden (2003) have presented auditory and acoustic analysis of controlled, read speech by four speakers of Standard Dutch, focussing on /r/'s segmental status and non-segmental correlates. This paper presents results of an attempt to replicate Plug & Ogden's study using uncontrolled and unscripted speech by 20 speakers. One might expect that in spontaneous speech, 'deletion' of /r/ is relatively common. The main question addressed in this paper is whether the non-segmental correlates of /r/ described by Plug & Ogden (2003) are as robustly observed in spontaneous speech as they are in laboratory speech. The findings suggest that this is not the case, although a subset of Plug & Ogden's correlates is clearly observable.

1. Introduction

This paper reports on a study of the realisation of postvocalic /r/ in stressed syllables followed by /t/ or /d/ in Dutch. The phoneme /r/ has received a great deal of attention in the literature on Dutch, because its realisation is highly variable across the Dutch language area, with a great deal of inter and intra-speaker variation (Van de Velde 1994, Van Reenen 1994, Van de Velde & Van Hout 1999). Van de Velde & Van Hout (1999) show that even in the standard variety of Dutch, a substantial number of realisations of /r/ can be observed, and many speakers use more than one realisation. Some of this variation is phonologically constrained: for example, the number of realisations used in syllable-initial position is much larger than that used in postvocalic position. With regard to postvocalic /r/, two further observations have been made. The first is that /r/ can be 'deleted' in this context (Cucchiarini & Van den Heuvel 1995, Collins & Mees 1996, Ernestus 2000, Van den Heuvel & Cucchiarini 2001, Verstraeten & Van de Velde 2001). The second is that /r/ affects the quality and duration of a preceding vowel and a following alveolar consonant (Koopmans-van Beinum 1969, Van den Heuvel et al. 1994, Clark & Yallop 1995, Collins & Mees 1996). It is these two observations that are the subject of investigation in this paper.

The phenomenon of '/r/-deletion' was subjected to detailed phonetic and phonological analysis by Plug & Ogden (2003) and Plug (2003). On the basis of a study of the speech of four Standard Dutch speakers, Plug & Ogden (2003) confirm that the realisation of postvocalic /r/ is variable, with segmental realisations including postalveolar and retroflex approximants; and that a segmental realisation is frequently impressionistically absent. They also confirm that /r/ has a number of recurrent non-segmental correlates, which can be observed even in tokens that lack a segmental realisation. They report that relative to tokens without /r/, tokens with /r/ have a longer vowel portion; a more central vowel quality and/or a schwa-like off-glide; a shorter final plosive burst portion; a greater proportion of gradual plosive releases; and a lower-frequency resonance throughout the plosive burst. They back up these observations through temporal, formant and spectral balance analyses, among others, showing that most of the correlates are statistically robust across tokens and speakers. On the basis of the same investigation, Plug (2003) argues that while '/r/-deletion' may be a useful label to refer to the absence of a segmental realisation of /r/, the robust presence of non-segmental correlates means that the phonological unit /r/ is *not* 'absent' or 'deleted', in the sense that the phonological contrasts in which /r/ participates are not neutralised.

A potential weakness in this argument is that the phonetic investigation that Plug & Ogden report on relies on word lists read out in a carefully controlled environment by a very small number of speakers. Most studies which have attempted to quantify the effects of '/r/-deletion', on the other hand, use spontaneous speech as their primary data (e.g. Cucchiarini & Van den Heuvel 1998, 1999, Kessens et al. 1999, Ernestus 2000, Van den Heuvel & Cucchiarini 2001). It may well be that the non-segmental correlates identified by Plug & Ogden (2003) are much less robust, or even absent, in spontaneous speech, and that neutralisation of contrasts involving /r/ – that is, /r/-deletion *proper* – is more common than Plug & Ogden's findings suggest. This paper reports on a study that addresses the first of these issues. The study repeated the measurements reported on by Plug & Ogden using relevant tokens from a corpus of spontaneous speech produced by 20 speakers of Standard Netherlands Dutch, in order to ascertain to what extent Plug & Ogden's findings can be generalised to spontaneous speech.

The remainder of the paper is organised as follows. Section 2 describes the corpus of speech used in this study, and the method of sampling relevant tokens from that corpus. Section 3 reports on an auditory analysis of a subset of tokens, which provides a first impression of the generalisability of Plug & Ogden's previous findings. Sections 4 and 5 report on a range of measurements taken across the entire data set: section 4 outlines the methods used, and section 5 summarises the results. Section 6 offers discussion and concludes the paper.

2. Data

The corpus used for this study is the Ernestus corpus of spontaneous Dutch conversation (Ernestus 2000). Ernestus' corpus contains speech by ten pairs of male speakers of Standard Dutch from the Western provinces of the Netherlands, mostly pairs of friends or colleagues, involved in several tasks, recorded in a professional recording studio. Most of the material comprises informal interviews which Ernestus undertook with each of the pairs (in which Ernestus asked a small number of questions designed to generate talk between the participants), and direct one-to-one conversations between the two members of each pair on a range of topics – some suggested by Ernestus, others offered spontaneously. In total, the material amounts to approximately 15 hours of spontaneous conversation.

Words ending in /rt/ and /rd/ and corresponding words ending in /t/ and /d/ were exhaustively sampled from the corpus, resulting in a data set of 372 tokens with /r/ (henceforth '+r tokens'), and 263 tokens without /r/ (henceforth '–r' tokens). Like Plug & Ogden (2003), the present study considers both monosyllabic words and polysyllabic words whose final syllable is stressed and contains the crucial segments. Function words were not included, unless produced with a marked sentence accent. The distribution of tokens across speakers is unbalanced, as one might expect. The number of +r tokens per speaker ranges between 4 and 35; the mean number of –r tokens between 3 and 29. The distribution of tokens is also unbalanced with regards to vowel quality. In Dutch, syllables ending in /rt/ or /rd/ can contain any of 12 vowels, which can be divided into three classes based on their phonological and phonetic behaviour: LONG /e ø o a/, SHORT /ɪ ɛ ɑ ɔ ʏ/, and HIGH /i y u/ (see Booij 1995, Van Oostendorp 2000). Long vowels can occur in open syllables, while short vowels cannot. The vowels in these two groups also differ in duration, as suggested by the labels. High vowels behave phonologically like long vowels, in that they can occur in open syllables. However, they are like short vowels in terms of duration. There was no fixed sample size for individual vowels or vowel types. Table 1 shows the distribution of +r and –r tokens across the three vowel types. For two vowels, no comparison between +r and –r contexts was possible: the corpus contains no +r tokens with /ɪ/, and no –r tokens with /ø/.

Table 1. Distribution of +r and –r tokens according to vowel type: LONG /e ø o a/, SHORT /ɪ ɛ ɑ ɔ ʏ/, HIGH /i y u/

	LONG	SHORT	HIGH	total
+r	237	83	52	372
–r	95	84	84	263
total	332	167	136	635

3. Auditory analysis

As indicated above, Plug & Ogden (2003) report that relative to tokens without /r/, tokens with /r/ have a longer vowel portion; a more central vowel quality and/or a schwa-like off-glide; a shorter final plosive burst portion; a greater proportion of gradual plosive releases; and a lower-frequency resonance throughout the plosive burst – while frequently lacking a segmental realisation of /r/. They present phonetic transcriptions of selected tokens that illustrate these features. As a first attempt to assess the extent to which Plug & Ogden's findings are generalisable to spontaneous speech, selected tokens from the present data set were transcribed on the basis of auditory analysis. Two phoneticians, including the author, independently produced impressionistic transcriptions for a sample of 10 +r and 10 –r tokens. They then compared transcriptions and agreed on an often slightly broader consensus transcription for each token.

Table 2 shows the consensus transcriptions. Segmental realisations of /r/ are not observed in this sample. In fact, auditory observation by the author alone suggests that segmental realisations of /r/ are very rare in the present data set – rarer than in Plug & Ogden's data set, which contained a substantial minority of tokens with a post-alveolar approximant realisation. On the other hand, some of the non-segmental correlates observed by Plug & Ogden are marked by both transcribers. Long vowels and schwa offglides are observed in +r tokens, but not in –r tokens; and vowels are more often marked as non-peripheral in +r tokens. Retracted coronal plosives are observed in +r tokens, too, although they are also transcribed in –r tokens. Affrication and incomplete closure are also observed in both +r and –r tokens. This suggests that several non-segmental correlates of /r/ described by Plug & Ogden (2003), specifically those relating to the final plosive burst, are less consistently observed in the present data set, but correlates relating to vowel duration and quality are observed across tokens. The remainder of the paper reports on the quantitative acoustic analysis of the correlates.

Table 2. Consensus transcriptions of selected –r tokens (left) and +r tokens (right)

–r token	transcription	+r token	transcription
niet	nitˢ	*viert*	fḭə̆t
instituut	ɪnsity̠t̞	*buurt*	by̠ə̆ɾ̞tˢ
meet	meɪtˢ	*garandeert*	xɐɾəndɪːt
pet	pɛd	*werd*	ʋɛə̆t̠ʰ
bestaat	bstatˢ	*kaart*	ka̠ːt̞
gat	χˣɑ̟tˢ	*apart*	əpʰɐɪ̆tˢ
bod	bo̞t̞	*kort*	ko̞ət
dood	d̥oʊt	*soort*	s̠ɔ̝t
doet	d̥ut̠̞ʷ	*toert*	tuət̠
flut	flɤt	*gebeurd*	xəbɜ˞ːt̠

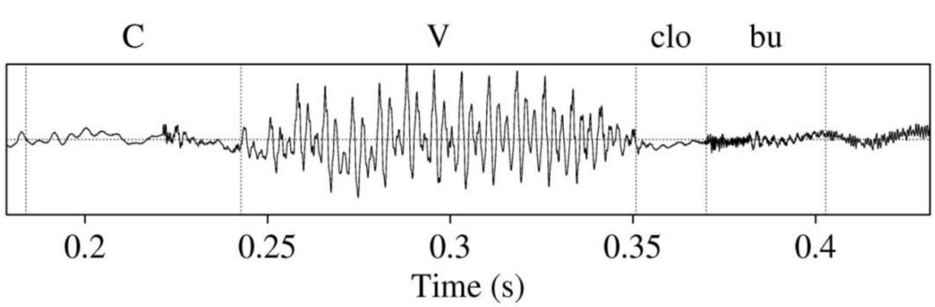
4. Acoustic analysis: Method

This section outlines the methods used in the quantitative analysis of the non-segmental correlates of /r/ described above. Following Plug & Ogden (2003), this study incorporates a range of analyses of +r and –r tokens: temporal analysis (section 4.1), formant analysis of vocalic portions (section 4.2), spectral balance analysis of plosive bursts (section 4.3), and analysis of the manner of plosive release (Section 4.4). Section 4.5 outlines the statistical tests applied in the analyses.

4.1 Segmentation and temporal analysis

Following Plug & Ogden (2003), tokens were segmented into an initial consonant portion (*C*); a vocalic portion (*V*); a stop closure portion (*clo*); and a burst portion (*bu*), to facilitate acoustic analysis. 46 tokens contain an unreleased final plosive; for these no burst portion could be segmented. Segmentation was done in PRAAT, and followed the general segmentation guidelines set out in Van de Heuvel et al. (1994) and Rietveld & Van Heuven (1997) for Dutch. Figure 1 shows one segmented token for illustration.

Figure 1. Segmentation of tokens; token: *kaart* /kart/



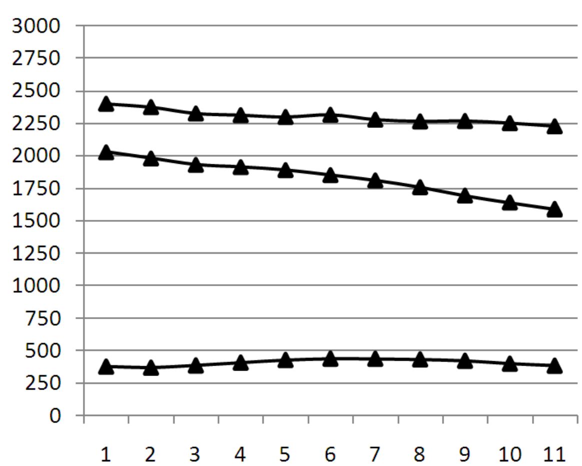
Durations of the segments were measured, and the crucial durations (those of *V* and *bu*) were expressed as a proportion of overall syllable duration as a rough control for speech rate variation. Normalisation for number of segments in the syllable Onset was not deemed necessary, since the +r and –r groups have similar proportions of tokens with one Onset segment only (88% and 81% respectively).

4.2 Formant analysis

For the purpose of formant analysis, measurements of F1, F2 and F3 were taken at 10% intervals throughout the vocalic portion (*V*), using the Burg LPC algorithm implemented in PRAAT, to produce stylised formant tracks such as that in Figure 2. The algorithm was run with default settings in the first instance. Where formant tracks were incompatible with a visual interpretation of the corresponding spectrogram, for example because two visually distinct formants were treated as one by the algorithm, the algorithm was run again with the number of assumed formants within the 0–5000Hz range adjusted, from default 5 to either 4 or 6. In most cases this resulted in an improved fit between the output of the algorithm and the visual interpretation of the spectrogram, and the output of the algorithm was accepted.

In a minority of cases, individual measurement points (for example, the F1 measurement at 10%, or the F3 measurement at 60%) clearly did not fit the overall contour of the formant tracks, because of a local failure of the algorithm to identify the formants accurately. These measurement points were adjusted manually to fit the overall contour. Formant tracks for which a clear contour could not be established, either by algorithm or by visual interpretation of the spectrogram, were excluded from further analysis. This left 591 tokens for the analysis of F1 and F2, and 546 tokens for that of F3.

Figure 2. Formant measurements taken at 10% intervals throughout the vocalic portion: F1 (bottom), F2 (middle) and F3 (top); vowel represented: /i/ in *viert* /virt/



4.3 Spectral balance analysis

For the purpose of spectral balance analysis of the plosive burst, measurements of the spectral centre of gravity (CoG) were taken at 25% intervals of the burst portion (*bu*). The spectral centre of gravity can be defined as the mean frequency-amplitude pairing of the spectrum (Van Son & Pols 1999, Jones & McDougall 2009). CoG measurements were taken with default settings in PRAAT in the first instance; however, a 1500–5000Hz pass Hann filter was applied when it became clear that F2 resonances persisting from the vowel portion lowered CoG values to well below the critical range reported by Plug & Ogden (2003) for a proportion of measuring points, making the overall range of values unrealistically wide both within and across tokens. With filtering, values were on the whole higher than those reported by Plug & Ogden, who did not apply a filter, but highly consistent within tokens. For each token, a mean CoG value was then computed.

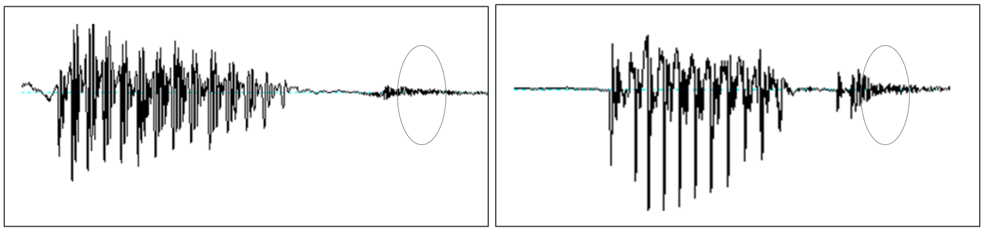
4.4 Burst classification

Following Plug & Ogden (2003), all final plosive releases were categorised as 'gradual' or 'abrupt' on the basis of visual inspection of the waveform: gradual releases are characterised by aperiodicity which gradually increases in amplitude, while abrupt releases are characterised by aperiodicity whose onset is marked by a spike in the waveform. Figure 3 illustrates the difference. In articulatory terms, the distinction relates at least partly to the speed of the plosive release: the slower the opening gesture, the more likely that in acoustic terms, the release is characterised by low-amplitude turbulence in the first instance.

4.5 Statistical analysis

Group effects with respect to continuous variables were tested through analysis of variance, with rhoticity (+r *vs* –r) and, where relevant, vowel type (long *vs* short *vs* high) as fixed factors and speaker identity as a random factor. Where relevant, *post-hoc* analysis was conducted by calculating Tukey's honestly significant difference. Any within-group effects of rhoticity were tested using independent sample *t*-tests, accompanied by Levene's test for equality of variance. Any correlations were evaluated using Pearson's correlation coefficient. Group effects for frequency variables were tested using Pearson's *chi*-square and Fisher's exact test. In cases of agreement, only the former is reported.

Figure 3. Waveforms illustrating the difference between a gradual (left) and abrupt (right) plosive release; crucial portions circled



5. Acoustic analysis: Results

This section summarises the results of the quantitative analysis of the non-segmental correlates of /r/ described above. The order of presentation roughly follows that of the previous section: section 5.1 presents the main results of the temporal analysis, section 5.2 presents the main results of the formant analysis, and section 5.3 presents the main results of the spectral balance and burst classification analyses.

5.1 Temporal analysis

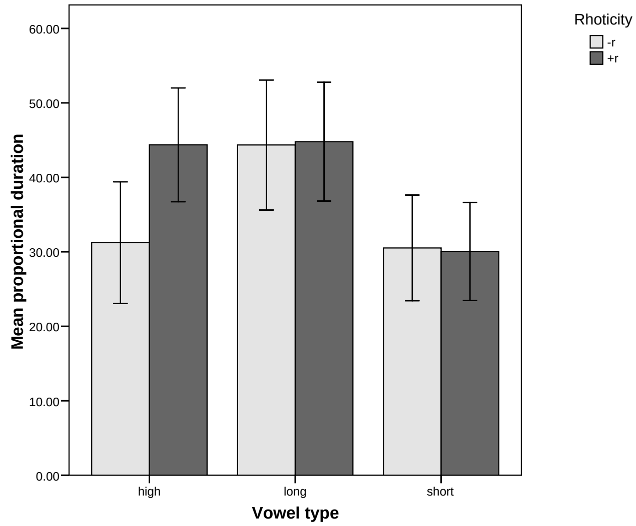
As indicated in the introduction, Plug & Ogden (2003) report that relative to –r tokens, +r tokens have long vowel portions and short burst portions. Statistical analysis of the duration measurements described above reveals, first of all, that the vowel lengthening effect of /r/ observed by Plug & Ogden (2003), as well as Nooteboom (1972), Van den Heuvel (1996) and Collins & Mees (1996), is observed across the present data set, too. On average, +r vocalic portions constitute 41.18% of the overall token duration, while –r tokens constitute 35.63%. This difference is highly significant ($F(1)=34.42$, $p<0.001$). As indicated above, +r and –r groups have similar proportions of tokens with one Onset segment only. In fact, there is no significant correlation between the proportional duration of the vocalic portion and raw token duration ($R^2=0.002$, $p=0.29$). This means it is unlikely that the difference between +r and –r tokens in the proportional duration of the vocalic portion can be attributed to differences in overall duration between +r and –r tokens, for example due to differences in the number of Onset segments. In addition, significant effects are observed for the fixed factor vowel type ($F(2)=108.26$, $p<0.001$), as well as the interaction between rhoticity and vowel type ($F(2)=22.59$, $p<0.001$). The random factor speaker and all other interactions are non-significant.

The significant interaction between rhoticity and vowel type in conditioning the proportional duration of the vocalic portion warrants a closer look at the effect of rhoticity within the subgroups of high, long and short vowels. This reveals that the lengthening effect of /r/ is only observed for the high vowels /i y u/, as illustrated in Figure 4. Within this subset, which contains 52 +r tokens and 84 –r tokens, +r vocalic portions have a significantly higher variance (*F*=19.03, *p*<0.001) and mean (*t*(73.45)=10.61, p<0.001 assuming unequal variance) than –r vocalic portions. Figure 4 shows that in the +r context, high vowels are, on average, similar in duration to long vowels, while in the –r context they are similar to short vowels. This patterning has been observed before (Nooteboom 1972), so it is not unexpected that the effect of rhoticity is *largest* for the high vowels. However, Plug & Ogden (2003) report a significant effect of rhoticity for all vowel types. In the present data set, by comparison, long and short vowels are unaffected by rhoticity, and the significant main effect of rhoticity across all vowels is due only to its effect on high vowel duration.

Turning now to the burst duration, as indicated above 46 tokens lack a delimitable burst portion. These were excluded from the burst duration analysis. In the remaining set of 589 tokens, the first thing to observe is that there is a significant negative correlation between the proportional duration of the burst and the proportional duration of the vowel (R^2=0.03, *p*<0.001). In other words, the greater the proportional duration of the vowel, the smaller that of the plosive burst. However, there is also a significant *positive* correlation between the proportional duration of the burst and the raw token duration (R^2=0.02, *p*<0.001), which means that it is not the case that tokens which are relatively long due to a long vocalic portion invariably have relatively short proportional burst durations. Rather, it may be that a relatively low local articulation rate manifests itself particularly in the relative duration of the plosive burst.

With the two correlating factors covaried out, no significant effect of rhoticity is observed (*F*(1)=1.96, *p*=0.17). Interestingly, vowel type turns out to be a significant factor (*F*(2)=8.33, *p*<0.001), while the interaction between vowel type and rhoticity is non-significant. *Post-hoc* analysis suggests that there are significant differences in proportional burst duration between tokens with long and short vowels (*p*<0.001), and long and high vowels (*p*=0.001). This is likely to be indicative again of a correlation between vowel duration and burst duration, which is not further explored here. Importantly, rhoticity appears to have no impact on this correlation other than its impact on high vowel duration.

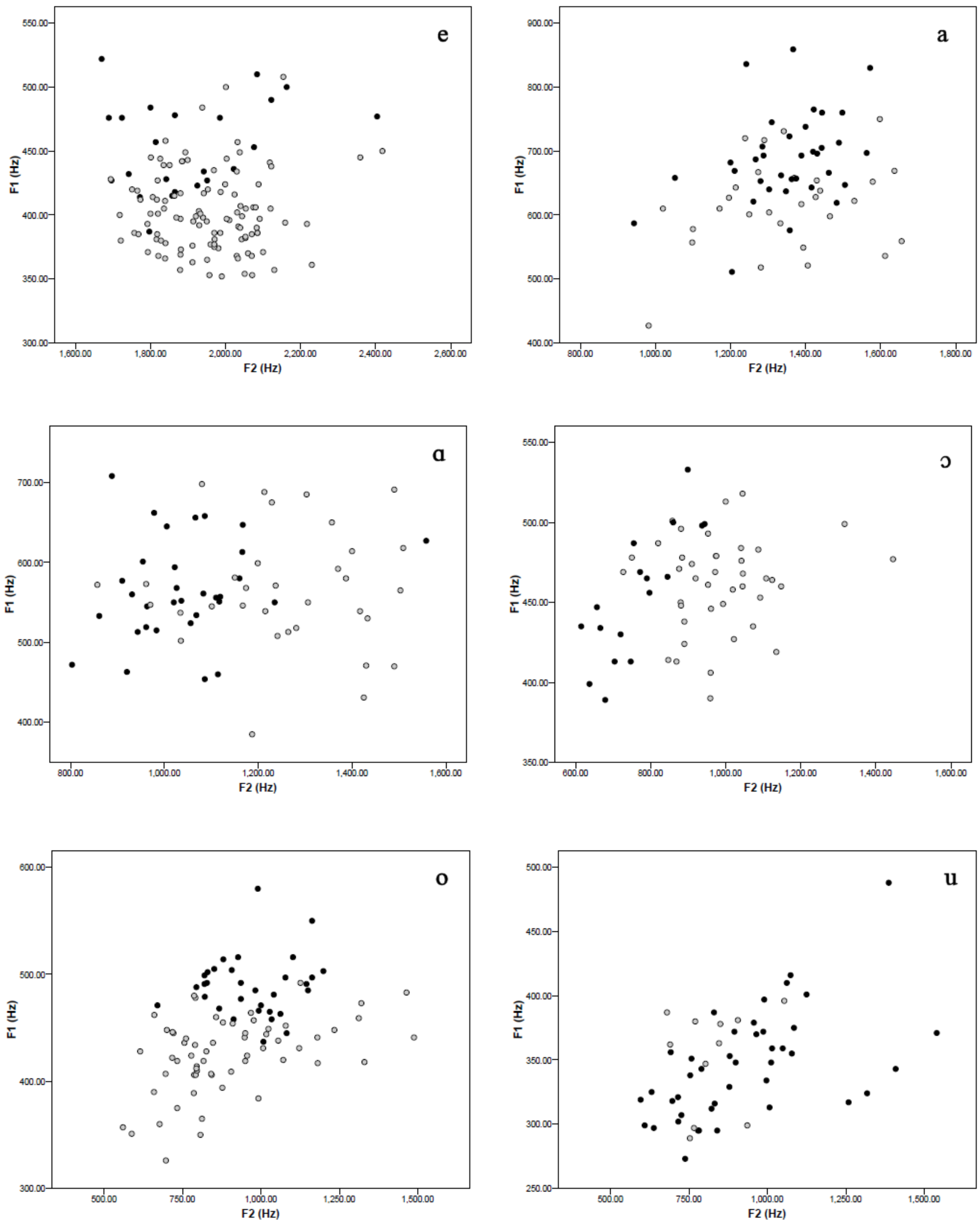
Figure 4. Effect of /r/ on vowel duration according to vowel type: LONG /e ø o a/, SHORT /ɪ ɛ ɑ ɔ ʏ/, HIGH /i y u/



5.2 Formant analysis

Plug & Ogden (2003) observe that relative to –r tokens, +r tokens have a more central vowel quality and/or a schwa-like off-glide. They report on measurements of F1, F2 and F3 at the midpoint and in the final periods of vocal-fold vibration, and show a relative reduction of the F1-F2 vowel space in +r tokens, as well as a lowering of F3 (see also Koopmans-van Beinum 1969, Schouten & Pols 1979). As indicated above, in the present investigation stylised formant tracks were created for each token by taking formant measurements at 10% intervals of the vocalic portion, and selected points were subjected to statistical analysis. Starting with F1 and F2 at the midpoint (measuring point 6 in Figure 2 above), evaluating the difference between +r and –r subsets of tokens was impossible for the vowels /ɪ/ and /ø/, since the data set contains tokens in one subset only, and not feasible for /i/, /ʏ/ and /ɛ/, since the data set does not contain enough tokens in either of the subsets. This leaves seven vowels. For each, measurements were plotted on an F1–F2 scattergram, and any differences between +r and –r tokens along either dimension tested statistically. The scattergrams are presented in Figure 5.

Figure 5. F1-F2 values for the vowels /e/, /a/, /ɑ/, /ɔ/, /o/ and /u/ in +r tokens (grey circles) and –r tokens (black circles), taken at the midpoint of the vocalic portion



A first thing to notice when looking across the scattergrams in Figure 5 is that for all seven vowels, the clouds of F1-F2 values in +r and –r tokens overlap to a considerable extent. In other words, the effect of /r/ on a preceding vowel is not a categorical quality change. Still, a statistically significant effect of rhoticity is observed for all vowels except /y/. Starting with /e/, rhoticity has a lowering effect on F1: the mean across +r tokens (N=106) is 402 Hz, while that across –r tokens (N=23) is 454 Hz ($F(1)= 33.28$, $p<0.001$). No significant difference between speakers is observed. For /a/, rhoticity also has a lowering effect on F1: the mean across +r tokens (N=29) is 615 Hz, while that across –r tokens (N=35) is 688 Hz ($F(1)= 10.26$, $p=0.006$).

Again, no significant difference between speakers is observed. For /ɑ/, rhoticity has a raising effect on F2: the mean across +r tokens (N=33) is 1255 Hz, while that across –r tokens (N=32) is 1044 Hz ($F(1)= 15.75$, $p<0.001$). A significant difference between speakers is observed ($F(17)= 2.63$, $p=0.048$), but the interaction between factors is not significant. For /ɔ/, rhoticity also has a raising effect on F2: the mean across +r tokens (N=41) is 985 Hz, while that across –r tokens (N=17) is 766 Hz ($F(1)= 14.98$, $p=0.003$). No significant difference between speakers is observed.

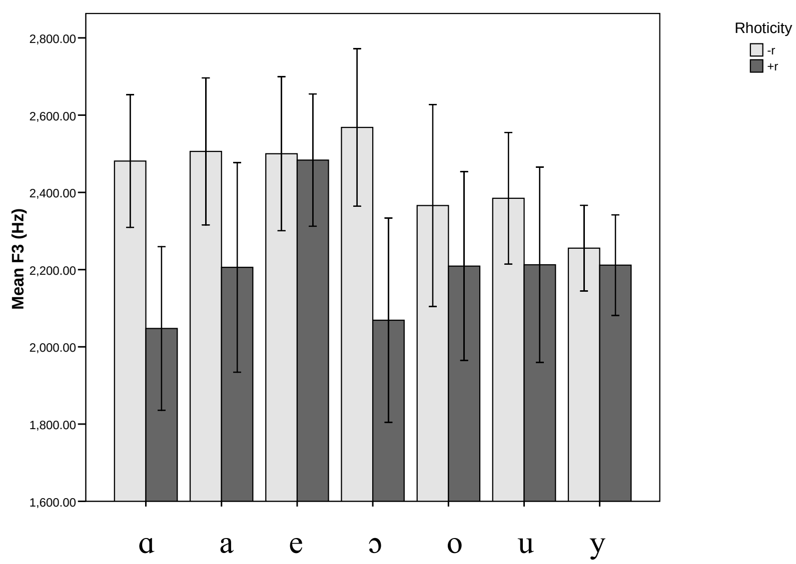
The effects for /o/ and /u/ are somewhat more complex. For /o/, rhoticity has a lowering effect on F1: the mean across +r tokens (N=61) is 425 Hz, while that across –r tokens (N=32) is 489 Hz ($F(1)= 34.31$, $p<0.001$). No significant difference between speakers is observed. Rhoticity also has an effect on F2: as seen in Figure 5, the range of F2 values is noticeably higher in +r tokens than in –r tokens (Levene's, $F=5.74$, $p=0.019$). No significant difference in mean is observed: that is, F2 values are more variable in +r tokens, around a similar mean as that observed for –r tokens. For /u/, rhoticity has a lowering effect on F2: the mean across +r tokens (N=11) is 823 Hz, while that across –r tokens (N=42) is 924 Hz. However, Figure 5 shows that the subset of –r tokens is not homogeneous, containing a small minority of tokens with F2 values above 1250 Hz. As a result, the range of F2 values is noticeably higher in –r tokens than in +r tokens (Levene's, $F=4.88$, $p=0.032$). Removing the outliers and running a regular analysis of variance results in the observation of a significant effect of rhoticity; the same result is obtained from a *t*-test assuming unequal variances ($t(33.56)= –2.09$, $p=0.044$).

To summarise, /r/ has no significant effect on /y/; a lowering effect on F1 for /e/, /a/ and /o/; a raising effect on F2 for /ɑ/ and /o/; and a lowering effect on F2 for /u/. While these effects do not constitute the overall mid-centralisation that Plug & Ogden (2003) and others (Koopmans-van Beinum 1969, Van Oostendorp 2000) refer to, they are to some extent consistent with a compression of the vowel space in +r tokens: in articulatory terms, front open vowels are less open (lowering effect on F1 for /e/ and /a/) and back vowels are less back (raising effect on F2 for /ɑ/ and /o/). The results for /u/ do not fit this pattern, however: the lowering effect on F2 suggests that /u/ is further back in +r tokens.

Differences between +r and –r tokens in the extent of liprounding may be relevant in the F2 dimension, but these must be left for future research.

Turning now to F3 at the midpoint, rhoticity has a significant lowering effect across the data set ($F(1)=43.35$, $p<0.001$). However, there are also significant differences in F3 values for +r and –r tokens across vowels (factor vowel: $F(11)=7.18$, $p<0.001$; interaction between rhoticity and vowel: $F(7)=6.94$, $p<0.027$), and across speakers (factor speaker: $F(19)=3.24$, $p<0.001$; interaction between rhoticity, vowel and speaker: $F(43)=1.87$, $p=0.001$). Figure 6 illustrates the differences across vowels. With speaker as a random factor, the effect of rhoticity is significant for /a/ ($F(1)=37.12$, $p<0.001$; no significant interaction with factor speaker), /ɑ/ ($F(1)=55.49$, $p<0.001$; no significant interaction with factor speaker), /o/ ($F(1)=14.18$, $p=0.001$; no significant interaction with factor speaker) and /ɔ/ ($F(1)=14.47$, $p=0.005$; significant interaction with factor speaker, $F(8)=9.28$, $p<0.001$). Speaker differences are not further examined here. For /e/, /y/ and /u/ no significant effect of rhoticity or the interaction between rhoticity and speaker is observed, although all three have a lower mean in +r tokens; while for /i/, /ʏ/ and /ɛ/ the subsets are too small to allow for statistical testing, as indicated above.

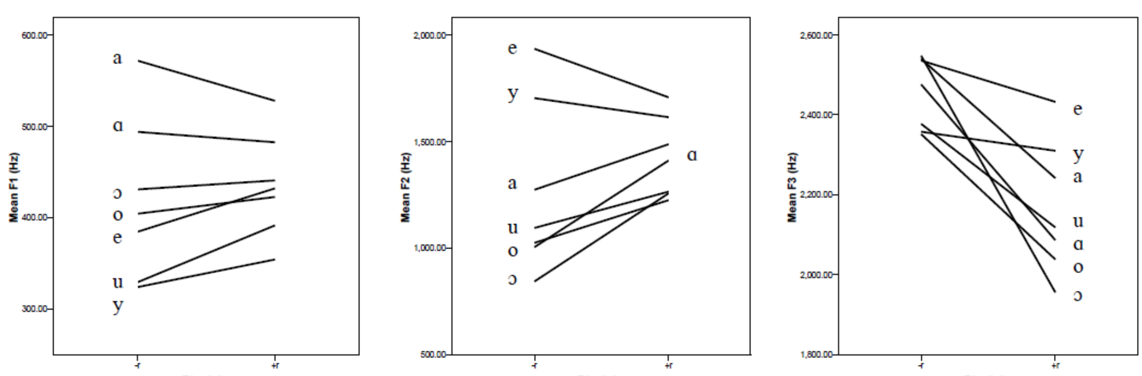
Figure 6. Mean F3 values (with SD) for the vowels /ɑ/, /a/, /e/, /ɔ/, /o/, /u/ and /y/ across +r tokens (dark grey) and –r tokens (light grey), taken at the midpoint of the vocalic portion



Finally, Plug & Ogden (2003) report F3 lowering and mid-centralisation across vowels and speakers during the final 28 ms of the vocalic portion. In the present investigation, measurements for F1 to F3 at the start of the final 10% of the vocalic portion (measurement point 10 in Figure 2 above) were used to test whether these effects are observed in the spontaneous speech data. Figure 7 summarises the data. Each line represents one of the seven vowels considered, and connects the mean formant values across –r tokens (left) and +r tokens (right). The figure shows that for F1 and F2, a degree of mid-centralisation is observed: for both formants, the range of formant values is smaller in +r tokens than in –r tokens, and the +r range is contained within the –r range.

Thus, vowels with relatively low F1 or F2 values –r tokens have higher corresponding values in +r tokens; vowels with relatively high values in –r tokens have lower corresponding values in +r tokens; and vowels whose F1 or F2 values are in the middle of the overall ranges of values show little difference between +r and –r tokens. Across vowels, the effect of rhoticity on F1 is not significant; however, the interaction between rhoticity and vowel is ($F(6)=2.75$, $p=0.021$). The random factor speaker has no significant effect. The effect on F2 is significant across vowels ($F(1)=30.11$, $p<0.001$), and again, a significant interaction between rhoticity and vowel is observed ($F(6)=11.35$, $p<0.001$). For F2, the random factor speaker also has a significant effect ($F(19)=2.43$, $p=0.040$), but no significant interaction between speaker and rhoticity is observed.

Figure 7. Mean F1 (left), F2 (middle) and F3 (right) values for the vowels /ɑ/, /a/, /e/, /ɔ/, /o/, /u/ and /y/ across –r tokens (left) and +r tokens (right), taken at the start of the final 10% of the vocalic portion



Looking at Figure 7, for F3 the effect of rhoticity is clearly one of lowering rather than centralisation: all vowels except /y/ have a considerably lower F3 in +r tokens than in –r tokens. The effect is highly significant across vowels ($F(1)=54.78$, $p<0.001$), although there is a significant interaction between rhoticity and vowel ($F(6)=4.56$, $p=0.001$). The random factor speaker has no significant effect. The lowering effect on F3 is significant for the individual vowels /a/ ($F(1)=14.98$, $p=0.003$; no significant interaction with factor speaker), /ɑ/ ($F(1)=51.02$, $p<0.001$; significant interaction with factor speaker, $F(12)=2.22$, $p=0.037$), /o/ ($F(1)=16.18$, $p=0.001$; no significant interaction with factor speaker) and /ɔ/ ($F(1)=27.09$, $p=0.001$; significant interaction with factor speaker, $F(8)=2.66$, $p=0.030$). The significant interactions with the factor speaker suggest that not all speakers show the lowering effect of /r/ to the same degree. For /e/, the interaction between rhoticity and speaker is significant ($F(12)=2.29$, $p=0.014$), but there is no significant main effect of rhoticity. For /y/ and /u/, no significant effects are observed. Notice that these results are almost identical to those for F3 at the midpoint of the vocalic portion. In other words, if a lowering effect of rhoticity is observed, as is the case for /a/, /ɑ/, /o/, it is observed at least from the midpoint onwards; the effect does not emerge during the latter half of the vocalic portion.

To summarise the results of the formant analysis, most of the differences between vowels in +r and –r tokens described by Plug & Ogden (2003) are observed in the present data set, too: we find evidence of mid-centralisation, F3 lowering and a centralising offglide. Most vowels for which detailed comparison was possible show effects of rhoticity on F1, F2 and F3 at the midpoint and in the final periods of vocal-fold vibration. The vowel /y/ is exceptional: it appears to be largely unaffected by /r/ in terms of quality, showing no significant difference between +r and –r tokens in F1, F2 and F3 values at the midpoint. The vowel /u/ also deserves further attention: while it shows an effect of /r/ on F2 values at the midpoint, the effect is not compatible with mid-centralisation.

5.3 Spectral balance analysis and burst classification

Plug & Ogden (2003) report that relative to –r tokens, +r tokens have a greater proportion of gradual plosive releases and a lower-frequency resonance throughout the plosive burst. On the basis of these observations, they hypothesise that postvocalic coronal plosives in +r tokens are apico-postalveolar, while those in –r tokens are lamino-alveolar. The relative retraction in +r tokens would explain the lower-frequency resonance, and apicality is associated with a relatively fast release gesture, which would explain both the greater proportion of gradual plosive releases and the shorter burst durations that characterise Plug & Ogden's +r tokens. Of course, we have already seen that the present data set does not show the effect of /r/ on plosive burst duration reported by Plug & Ogden.

Turning to the plosive burst resonance, statistical analysis of the mean centre of gravity of the plosive burst spectrum reveals that while the random factor speaker has a significant effect on values ($F(19)=5.65$, $p=0.027$), rhoticity does not have a significant effect. The factor vowel is also non-significant, and there are no significant interactions between rhoticity, vowel and speaker. Looking across speakers, we see that one speaker out of 20 only shows a significant lowering effect of rhoticity ($t(41)= -3.31$, $p=0.002$).

With respect to burst classification (gradual *vs* abrupt), significant effects of rhoticity ($\chi^2(1)=4.20$, $p=0.041$) and speaker ($\chi^2(19)=57.75$, $p<0.001$) are observed. The effect of rhoticity is rather limited, however: while in –r tokens, 44% of plosive releases are abrupt and 56% gradual (N=248), in +r tokens the balance is 53% *vs* 47%, respectively (N=339). In other words, the difference is one between a marginal minority (–r tokens) and a marginal majority (+r tokens) of abrupt releases. Note that on the whole, gradual releases are more common than abrupt releases in the present data set. This is not the case in Plug & Ogden's data set, and the difference is likely to be one of spontaneous *vs* elicited speech. Looking across speakers, we see that only two speakers show a significant effect in the overall direction (that is, +r tokens having more abrupt plosive releases than (–r tokens), while one speaker shows a significant effect in the opposite direction.

In sum, the effects of rhoticity on plosive place and manner of articulation reported by Plug & Ogden (2003) are only weakly observed in the present data set, with inter-speaker variation and, possibly, a greater tendency to plosive reduction reducing any effect that rhoticity might have.

6. Summary and outlook

This paper has reported on a study of the realisation of postvocalic /r/ in stressed syllables followed by /t/ or /d/ in Dutch. It set out to establish the extent to which the findings of Plug & Ogden (2003) with regards to non-segmental phonetic correlates of postvocalic /r/ are generalisable to spontaneous speech. The study suggests that only a subset of the correlates described by Plug & Ogden are robustly observable in spontaneous data. First, rhoticity has a significant effect on vowel duration – but only in tokens with the vowels /i y u/. Second, /r/ has a number of effects on the quality of a preceding vowel, as reflected in its F1, F2 and F3 values from the midpoint onwards. Still, this study has confirmed these effects – a reduction of the overall size of the F1-F2 vowel space and a lowering effect on F3 – for the vowels /e a ɑ ɔ o u/ only. Of the remaining vowels, /i ɪ ʏ ɛ ø/ could not be considered in detail, and /y/ was found to be largely unaffected by a following /r/. Moreover, it is worth repeating that the effect of /r/ on a preceding vowel does not appear to be a categorical quality change: while +r and –r tokens show different ranges of vowel qualities for /e a ɑ ɔ o u/, these ranges overlap to a considerable extent for each vowel. Third, a significant, albeit very weak effect of rhoticity on the likelihood of an abrupt (as opposed to a gradual) plosive release was observed. The remaining correlates reported by Plug & Ogden (2003), which can all be related to the place and manner of articulation of coronal plosives following /r/, were not found to be robust in the present data set.

The study reported here confirms, first of all, that laboratory speech and spontaneous speech are different, and that findings regarding speech patterns based on laboratory speech cannot be assumed without question to be generalisable to ordinary, everyday conversation (cf. Ernestus 2000). This is especially the case when the speech patterns under consideration involve phonetic reduction, as is the case here: it is arguably to be expected that spontaneous speech will display an overall higher degree of phonetic reduction than laboratory speech, and therefore that phonetic correlates observed in the latter may be absent in the former. In the present case, it appears that the effects of postvocalic rhoticity on the place and manner of articulation of coronal plosives displayed in more careful speech are absent in spontaneous speech. It may be worth noting that it is also possible that the apparent absence of these effects is due to coarticulation with following segments. In the present study, the nature of the sounds following the coronal plosive was not considered as a variable, while Plug & Ogden (2003) controlled for it by embedding all tokens in the same carrier phrase.

While a more careful analysis of the phonological environment of the coronal plosives in the present data set may confirm that in some contexts, the effects of rhoticity described by Plug & Ogden are observable, from a listener's point of view they hardly constitute a reliable cue for the presence or absence of a postvocalic /r/ (cf. Janse et al. 2007). Moreover, the significant effect of rhoticity on the likelihood of an abrupt plosive release is so weak that it is unlikely to be of much help in the classification of individual tokens. Still, the findings presented in this paper suggest that in the absence of a segmental realisation of /r/, a consideration of vowel quality and duration may in many cases result in an accurate classification of a token as +r or –r. Notice that a consideration of both is needed: the vowel for which no consistent quality effect of /r/ was observed, /y/, is one of the HIGH vowels, and can therefore be expected to show a robust effect of /r/ on vowel duration.

This brings us to the second issue raised in the introduction to this paper: if the non-segmental correlates identified by Plug & Ogden (2003) are much less robust, or even absent, in spontaneous speech, neutralisation of contrasts involving /r/ may be more common than Plug & Ogden's findings suggest. Further research is needed to establish how many tokens in the present data set have none of the correlates of postvocalic /r/ described here, and what proportion of tokens can be correctly classified as +r or –r on the basis of phonetic analysis alone. This will provide a valuable insight into the development of postvocalic rhoticity in Dutch, as well as into the nature of segmental 'deletion'.

Acknowledgements

I would like to thank Barry Heselwood for his contribution to this research, and Mirjam Ernestus for valuable discussion of an earlier version of this paper, presented at *Form and Function of Insertions and Deletions in Speech*, Osnabrück, March 2009. I am also grateful to Malcah Yaeger-Dror and an anonymous reviewer for critical comments on a draft of this paper.

References

Booij, G. E. (1995). *The Phonology of Dutch*. Oxford: Oxford University Press.

Clark, J. & Yallop, C. (1995). *An Introduction to Phonetics and Phonology*. Oxford: Blackwell. Second edition.

Collins, B. & Mees, I. M. (1996). *The Phonetics of English and Dutch*. Leiden: E. J. Brill. Second edition.

Cucchiarini, C. & Van den Heuvel, H. (1995). /r/-deletion in Standard Dutch. *Proceedings of the Department of Language and Speech, University of Nijmegen* **19**, 59–65.

Cucchiarini, C. & Van den Heuvel, H. (1998). Postvocalic /r/-deletion in standard Dutch: How phonological research can profit from ASR technology. *Proceedings International Conference on Spoken Language Processing*, Vol. 5, 1759–1763

Cucchiarini, C. & Van den Heuvel, H. (1999). /r/-deletion in Dutch: more experimental evidence. *Proceedings of ICPhS 1999*, Vol. 3, 1673–1676.

Ernestus, M. (2000). *Voice Assimilation and Segment Reduction in Casual Dutch: A Corpus-based Study on the Phonology–Phonetics Interface*. Netherlands Graduate School of Linguistics LOT, Utrecht.

Janse, E., Nooteboom, S. G., & Quené, H. (2007). Coping with gradient forms of /t/-deletion and lexical ambiguity in spoken word recognition. *Language and Cognitive Processes* **22**, 161–200.

Jones, M. J. & McDougall, K. (2009). The acoustic character of fricated /t/ in Australian English: A comparison with /s/ and /ʃ/. *Journal of the International Phonetic Association* **39**, 265–289.

Kessens, J. M., Wester, M. & Strik, H. (1999). Improving the performance of a Dutch CSR by modeling within-word and cross-word pronunciation variation. *Speech Communication* **29**, 193-207.

Koopmans-van Beinum, F. J. (1969). Nog meer fonetische zekerheden. *De Nieuwe Taalgids* **62**, 245–250.

Nooteboom, S. (1972). *Production and Perception of Vowel Duration: A Study of Durational Properties of Vowels in Dutch*. Ph.D. dissertation, Rijksuniversiteit Utrecht.

Plug, L. (2003) Zo spreek je de *r* dus niet uit: Een fonetische kijk op ‘deletie’. *Nederlandse Taalkunde* **8**, 2–13

Plug, L. & Ogden, R. (2003). A parametric approach to the phonetics of postvocalic /r/ in Dutch. *Phonetica* **60**, 159–186.

Rietveld, A. C. M. & Van Heuven, V. J. (1997). *Algemene Fonetiek*. Coutinho, Bussum.

Schouten, M. E. H. & Pols, L. C. W. (1979). Vowel segments in consonantal contexts: A spectral study of coarticulation – Part 1. *Journal of Phonetics* **7**, 1–23.

Van de Velde, H. (1994). 60 jaar (r)evolutie in het Standaard-Nederlands. *Taal en Tongval*, theme issue **7**, 22–42.

Van de Velde, H. & Van Hout, R. (1999). The pronunciation of (r) in Standard Dutch. In *Linguistics in the Netherlands 1999*. In van Bezooijen, R & Kager, R (eds.) 177–188. Amsterdam: John Benjamins.

Van den Heuvel, H. (1996). *Speaker Variability in Acoustic Properties of Dutch Phoneme Realisations*. Ph.D. dissertation, Universiteit van Nijmegen.

Van den Heuvel, H. & Cucchiarini, C. (2001). R-deletion in Dutch. Rumours or reality? *Etudes & Travaux* **4**, 185–198.

Van den Heuvel, H., Rietveld, A. C. M. & Cranen, B. (1994). Methodological aspects of segment- and speaker-related variability. *Journal of Phonetics* **22**, 386–406.

Van Oostendorp, M. (2000). *Phonological Projection: A Theory of Feature Content and Prosodic Structure*. Mouton De Gruyter, Berlin.

Van Reenen, P. (1994). Driemaal /r/ in de Nederlandse dialecten. *Taal en Tongval*, theme issue 7: 54–72.

Van Son, R. J. J. H. & Pols, L. C. W. (1999). An acoustic profile of consonant reduction. *Speech Communication* 28, 125–140.

Verstraeten, B. & Van de Velde, H. (2001). Socio-geographical variation of /r/ in Dutch. *Etudes & Travaux* 4, 45–61.

Leendert Plug
Dept of Linguistics & Phonetics
University of Leeds
LEEDS LS2 9JT
UK

l.plug@leeds.ac.uk