POSTVOCALIC /r/-DELETION IN STANDARD DUTCH: HOW EXPERIMENTAL PHONOLOGY CAN PROFIT FROM ASR TECHNOLOGY

Catia Cucchiarini and Henk van den Heuvel

A[2]RT, Dept. of Language and Speech, University of Nijmegen, the Netherlands
{catia, heuvel}@let.kun.nl, http://lands.let.kun.nl/

ABSTRACT

In this study automatic speech recognition (ASR) techniques were used to substantiate the findings of phonological research on postvocalic /r/-deletion in Standard Dutch. A database containing spontaneous speech utterances stemming from man-machine interactions in an automatic train-table inquiry system was used for this purpose. Pronunciation variants with and without /r/ were automatically generated on the basis of our specification of the phonological rule of /r/-deletion and were then included in the lexicon of a continuous speech recognizer (CSR) which was used in forced recognition mode. The results show that in a corpus containing 214,102 words, in which /r/-deletion could be applied 16,865 times, it was actually applied in 47.6% of the cases. This is a high percentage of occurrence for a phenomenon that has so sporadically been described. Furthermore, the results substantiate our rule specification: /r/-deletion occurs more often after schwa than after any other vowel.

1. INTRODUCTION

Rhotic sounds are known to be rather common in the languages of the world [1]. Although they behave in a similar way from a phonological point of view, phonetically, they appear to be rather different from each other [2]. Considerable variation in the realization of /r/-sounds can be found also within one and the same language. This is clearly the case in Dutch, where up to 10 different realizations have been distinguished for postvocalic /r/ alone [3: 130-131]. In general, the greatest amount of variation is observed when /r/ appears in coda position [2, 4, 5]. For Dutch it was also found that /r/-realizations in postvocalic position tend to exhibit a lower degree of articulatory strength, sometimes even resulting in complete deletion [3, 4].

This paper deals with /r/ in postvocalic position and before another consonant, since here /r/-deletion is most obvious in Dutch. In a previous paper [6], we accounted for the phenomena of postvocalic /r/-weakening and /r/-deletion in Dutch by referring to the linguistic context in which these processes are more likely to occur. On the basis of the data presented, it was concluded that in connected speech, an optional word phonology rule of /r/-weakening is operative in Dutch, which, in certain contexts, may lead to complete deletion of postvocalic /r/. /r/-weakening can be observed in all positions where /r/ is preceded by a vowel and followed by a consonant, whereas /r/-deletion does not occur everywhere. The decisive factor to be acknowledged is the left vowel context (see further section 2).

The best way to test any theory about these (and other) synchronic phenomena is to study them in real-life speech. However, in [6] we also pointed out some of the problems involved in studying connected speech processes. First of all, we underlined the difficulties involved in describing connected speech processes in categorical terms, (see also [7: 126]) because certain phenomena may exhibit a gradual nature, so that sometimes it can be very difficult to determine whether a segment is extremely weakened or completely deleted. Second, rules of connected speech are optional and their application is often influenced by factors such as speech style, speech rate, word frequency, socio-economic background and region of origin of the speaker, and, last but not least, individual variation. This implies that the application of these rules may be less amenable to observation according to phonological methods than that of obligatory rules. In other words, studying these kinds of phenomena would require transcribing considerable amounts of speech material, which is not always possible because transcribing speech is very time-consuming and therefore expensive. For this reason, in [6] we suggested the use of the large speech databases and the techniques that have been made available for the purpose of ASR in order to study these processes in real-life extemporaneous speech.

Following these considerations, in this paper we describe a study of postvocalic /r/-deletion in which a database containing spontaneous speech and a CSR were used to check whether the word-internal postvocalic /r/-deletion rule [6] takes place and in which contexts it occurs more often. Our aim for this paper is first to demonstrate that the optional word phonology rule of postvocalic /r/-deletion exists in Dutch, and, second, to show that it appears more often in some contexts than in others.

The paper is organized as follows. In section 2 the postvocalic /r/-deletion rule formulated in [6] is described. In section 3 the method adopted in the present study is presented. In particular, in 3.1 the speech database is shortly described. In 3.2 we explain how the different pronunciation variants were generated, and, finally, in 3.3 the technique of forced recognition is presented. In section 4 the results obtained with this technique are reported and in section 5 they are discussed in relation to those expected on the basis of the rule specification.

2. THE POSTVOCALIC /r/-DELETION RULE IN STANDARD DUTCH

In this section we will address postvocalic /r/-deletion in Standard Dutch as a function of the left vowel context of /r/, which, in our view, offers an appropriate framework to describe the phenomena in comprehensible phonological terms (see also, [6]).

2.1. Context 1: /r/-Deletion after Schwa

After schwa /r/-deletion is fully acceptable. Our intuitions are that in this context /r/-deletion is more likely than in the other contexts that will be described below (see [6]). Consider the following examples:

overmorgen	ov/ər/morgen	ov/ə/morgen
Rotterdam	Rott/ər/dam	Rott/ə/dam
middernacht	midd/ər/nacht	midd/ə/nacht
allerzielen	all/ər/zielen	all/ə/zielen
verbinding	v/ər/binding	v/ə/binding
uiterlijk	uit/ər/lijk	uit/ə/lijk

As can be seen, /r/-deletion applies before various types of consonants, both in prosodic compounds, that is combinations of two prosodic words, and within one prosodic word.

In the ultimate syllable /r/-deletion is also possible. In this position it will take place before the so-called appendix consonants /t/ and /s/, since, according to [7], no other consonant is allowed after schwa and /r/.

Albert	Alb/ər/t	Alb/ə/t
vaders	vad/ər/s	vad/ə/s

In summary, the rule for /r/-deletion after schwa can be formulated as follows:

```
r          ->          0 / ə _     X
                                   |
                                   +cons
```

2.2. Context 2: /r/-Deletion after Short Vowels

When considering /r/-deletion after short vowels, a distinction should be made between stressed and unstressed vowels. When the preceding short vowel is unstressed, /r/-deletion is possible.

parkeren	p/ɑr/keren	p/ɑ/keren
portier	p/ɔr/tier	p/ɔ/tier
hervormd	h/ɛr/vormd	h/ɛ/vormd

If /r/ is preceded by a short, stressed vowel, /r/-deletion would seem to be less probable.

Harlingen	H/ɑr/lingen	*H/ɑ/lingen
wortel	w/ɔr/tel	*w/ɔ/tel

However, in this paper we will confine ourselves to those cases in which /r/-deletion seems to be clearly allowed.

2.3. Context 3: /r/-Deletion after Long Vowels

The case of /r/ after a long vowel is special in more than one respect. Firstly, ensuing /r/ has the effect of significantly lengthening the duration of vowels such as /i,y,u/, which phonologically behave like long vowels, but are phonetically short [8]. Some degree of lengthening due to the presence of /r/ occurs also for the vowels /e,o,ø/. However, in this case the most obvious effect of the ensuing /r/ is a certain degree of diphthongisation towards schwa [9], which is also called 'centralisation'.

After a long vowel /r/-deletion is possible before all consonants, in prosodic words and in prosodic compounds, regardless of the stress pattern. However, in this context the centralisation caused by the presence of /r/ is retained even after /r/-deletion:

Haarlem	H/a^{ə}r/lem	H/a^{ə}/lem
noorman	n/o^{ə}r/man	n/o^{ə}/man
boerderij	b/u^{ə}r/derij	b/u^{ə}/derij

/r/-deletion also takes place in the ultimate syllable. In this case the rule can apply only before the appendices /t,s/, since, according to [10], no other consonant is allowed in this position after a long vowel followed by /r/.

paars	p/a^{ə}r/s	p/a^{ə}/s
poort	p/o^{ə}r/t	p/o^{ə}/t

3. METHOD

In order to check the correctness of the rule specification and to get an idea of the frequency with which postvocalic /r/-deletion occurs in spontaneous speech, an experiment was carried out in which an automatic CSR was forced to choose between alternative variants of one and the same word, which differed from each other in the presence/absence of /r/ in postvocalic position. In this experiment we set out to determine 1) whether /r/-less variants are selected by the CSR and 2) whether they occur more often when /r/ appears after schwa, as would be in line with our expectations.

3.1. The Speech Database

Since we wanted to investigate the phenomenon of /r/-deletion in real-life extemporaneous speech, a database was used that contains spontaneous speech recorded over telephone lines which stem from man-machine interactions in an automatic train time-table inquiry system [11]. The waveform format of the speech files is A-law. The database used in this investigation comprised 214,102 words in 73,343 utterances from 3,529 different dialogues. This is the training partition of the VIOS2 database. A total of 16,865 possible locations for /r/-deletion was found in the database. For each of the three types of vowel contexts specified in section 2 there were over 2,000 possible cases of /r/-deletion available (see further section 4).

3.2. Variant Generation

On the basis of the /r/-deletion rule as specified in 2, variants with /r/ and variants without /r/ were generated for all words in the database in which the relevant contexts were met. These variants were generated automatically by an implementation in Perl of the rule set described in 2. The total number of words concerned is 283. For each of these words the pronunciation variants with and without /r/ were included in the lexicon that was used for the forced recognition to be described in the following section.

3.3. Forced Recognition

In the present experiment a standard CSR was used with context-independent HMMs for 35 monophones. The models were trained on the canonical transcriptions of the words. This means that during training /r/ was not deleted anywhere from the phoneme transcriptions of the words. /l/ and /r/ had separate models for prevocalic and postvocalic position in the syllable. The postvocalic model of /r/ is tested for the /r/-deletion experiments reported on here. Each monophone consists of three segments of two equal HMM states. Speech is coded as 14 mel-based cepstra (c[0]-c[13]) and 14 corresponding delta cepstral coefficients. The frame width is 16 ms and frame shift is 10 ms. For details about the CSR, the reader is referred to [11].

The phoneme models were trained on the database described in 3.1. The fact that test and training database are the same is not considered a relevant disadvantage for forced recognition on a database of this size. Forced recognition entails that the CSR is forced to choose from among alternative variants of the same word. This procedure is usually applied when one is not interested in determining which word was spoken, but which pronunciation form of that word, as in this study. In order to be selected, the various pronunciation variants have to be included in the lexicon.

When forced recognition is employed to check the correctness of phonological observations based on intuitions, as in the present experiment, it is important to know whether this technique is reliable or not. In other words, we want to be sure that the choices of the CSR used in forced-recognition mode are in line with the choices a panel of listeners would make. To check this an experiment was carried out [12] which showed that the degree of agreement between the CSR and the reference transcription was .52 (Cohen's κ), while for the listeners this value varied between .81 and .41. This means that the CSR succeeds in selecting the correct pronunciation variant for the /r/-deletion rule in 75.6%, of the cases, while for human listeners this percentage varies between 74% and 91.3% .

4. RESULTS

As mentioned above, the database used for the present experiment contained 214,102 words. In all these words, the /r/-deletion rule as specified in 2 could be applied in 16,865 cases, while it was actually applied 8,019 times, that is in 47.6% of the cases. This is a high percentage of occurrence, especially if we consider that the phenomenon of /r/-deletion in Standard Dutch has been so sporadically attested in the literature. If we now break down the results for the various contexts, we obtain the picture shown in Table 1.

As is clear from Table 1, /r/-deletion is much more common after schwa than after the other vowels, which is in line with our expectations.

	possible	**applied**	**percentage**
after schwa	**10302**	**6826**	**66.3**
- word final	250	172	68.8
- other	10,052	6,654	66.2
after short vowel	**2091**	**450**	**21.5**
after long vowel	**4472**	**743**	**16.6**
- word final	1,638	153	9.3
- other	2,834	590	21.5

Table 1 Number of possible applications, number of applications and percentage of application of /r/-deletion in the three contexts.

It is well conceivable that the percentage of /r/-deletions found depends on the /r/-model used. Since the postvocalic /r/ model was trained on words where the /r/ itself was assumed to be present, but in fact apparently was not, this model is itself affected by the /r/-deletion phenomenon. Therefore, we carried out another forced recognition in which the model for *pre*vocalic /r/ was used to test /r/-deletion in postvocalic position (in Dutch prevocalic /r/ cannot be deleted). For this analysis we found 88.5% /r/-deletions after schwa, 53.5% after short vowels, and 45.8% after long vowels, which are even higher proportions of /r/-deletion than those reported in Table 1.

5. DISCUSSION AND CONCLUSIONS

On the basis of the data presented in the previous section, /r/-deletion would seem to be a rather common process in Standard Dutch. It is therefore surprising that this phenomenon has been so little described in the literature so far. There may be several reasons for this. First, the fact that the Dutch /r/ is still in the process of changing [3] and the possibility that /r/-deletion is only a recent phenomenon. Second, the fact that studying connected speech processes in a comprehensive way is very difficult, because they are subject to a great deal of variation due to many different factors. In general, these processes are best studied by using sociolinguistic approaches in which all relevant factors are either systematically varied or kept under control by adopting a factorial design. Usually, the analyses used in this type of approach are so laborious, that the amount of material has to be limited to make the analyses feasible, with obvious consequences for the generalizability of the results. For this reason in this paper an automatic method was proposed, which makes it possible to study huge amounts of spontaneous speech in relatively short time.

Furthermore, interesting results can also be obtained with respect to the frequency of the various phenomena in the various contexts. As a matter of fact, the results of the present experiment can be used to substantiate our intuitions as to the occurrence of /r/-deletion in Standard Dutch. First of all, the results show that /r/-deletion is a phenomenon that should be reckoned with in Standard Dutch. Second, these results also confirm our intuitions as to the context in which /r/-deletion is more likely to occur: /r/-deletion turns out to be much more common after schwa than after any other vowel. In future analyses we will examine the effect of word frequency and word type on the proportion of /r/-deletions. Also the effect of other types of left and right contexts on /r/-deletion will be studied.

As to our method as such, we showed that the /r/-deletion results observed are affected by the type of HMM models chosen. In our view the use of the postvocalic /r/ model is the best choice to test /r/-deletion in postvocalic position. First, also human listeners are likely to mentally refer to an /r/ in *post*vocalic context to assess /r/-deletion in this position. And second, the /r/-deletion percentages found for the postvocalic /r/ model correspond more closely to the percentages observed for the human listeners. But if one considers the prevocalic /r/ models more realistic, then the percentages for postvocalic /r/-deletion that we presented in Table 1 appear to be very conservative estimates!

Quite clearly, the analysis presented in this paper does not produce such detailed information as a sociolinguistic study in which all influential factors are systematically varied or at least kept under control. However, it does provide reliable results as to the incidence of a given phenomenon in the context of a specific realistic application.

More insight into specific phonological processes could be obtained if some information concerning speech and speaker background variables were included in the database. For Dutch such information regarding sex, age and regional background of speakers is present in the Polyphone corpus [13]. An important drawback of this corpus is that it contains read rather than extemporaneous speech. The Corpus of (spontaneously) Spoken Dutch which will be recorded in the near future will doubtlessly be able to fill up this gap in the collection of Dutch SLR.

To conclude, it is obvious that recent developments in speech technology open up new possibilities for phonetic and phonological research because the observations of phoneticians and phonologists can easily be tested on large amounts of data. At least the study of /r/-deletion in Dutch made a significant step forward by applying the proposed method.

6. ACKNOWLEDGMENTS

We would like to thank Mirjam Wester and Judith Kessens for carrying out the forced recognition and making the results available to us. We thank Lou Boves for thoroughly reviewing an earlier version of this paper.

7. REFERENCES

1. Maddieson, I. *Patterns of sounds*, Cambridge University Press, Cambridge, 1984.

2. Lindau, M. “The story of /r/”, in Fromkin, V. (ed.) *Phonetic Linguistics: Essays in honor of Peter Ladefoged*, Academic Press, Orlando, 157-168, 1985.

3. Van de Velde, H. *Variatie en verandering in het gesproken Nederlands (1935-1993)*, PhD Thesis, University of Nijmegen, 1996.

4. Vieregge, W.H. and Broeders, T. “Intra- en interspeaker variation of /r/ in Dutch”, *Proc. Eurospeech93*, Berlin, 267-270, 1993.

5. Van Reenen, P. “Driemaal -r- in de Nederlandse dialecten”, *Taal en Tongval* 7, 54-72, 1994.

6. Cucchiarini, C. and Van den Heuvel, H. “/r/-deletion in Standard Dutch”, *Proc. Department of language and Speech* 19, 59-65, 1995.

7. Booij, G. *The phonology of Dutch*, Clarendon Press, Oxford, 1995.

8. Gussenhoven, C. and Broeders, T., *The pronunciation of English. A course for Dutch learners*. Wolters- Noordhof-Longman , Groningen, 1976.

9. Koopmans-van Beinum, F. “Nog meer fonetische zeker-heden”, *De Nieuwe Taalgids* 62, 245-250, 1969.

10. Kager, R.W.J., *A metrical theory of stress and distressing in English and Dutch*, PhD Thesis, University of Utrecht, 1989.

11. Strik, H., Russel, A., Van den Heuvel, H., Cucchiarini, C., Boves, L., “A spoken dialogue system for the Dutch public transport information service”, *International Journal of Speech Technology*, 121-13, 1997.

12. Kessens, J., Wester, M., Cucchiarini, C., and Strik, H., “The selection of pronunciation variants in spontaneous speech: Comparing the performance of man and machine”, *Proc. ICSLP98*.

13. Den Os, E.A., Boogaart, T.I, Boves, L., and Klabbers, E. “The Dutch Polyphone Corpus”. *Proc. Eurospeech95*, Madrid,Vol I, 825-828, 1995.