/r/-deletion in Dutch: rumours or reality?*

Henk van den Heuvel & Catia Cucchiarini

A²RT, Dept. of Language and Speech, University of Nijmegen, the Netherlands
{H.v.d.Heuvel, C.Cucchiarini}@let.kun.nl, http://lands.let.kun.nl/

ABSTRACT

In this study we investigate three properties of the left vowel context which we hypothesize to underlie deletion of /r/ in postvocalic, preconsonantal position in Dutch spontaneous speech: vowel type (schwa, full vowel), vowel length (long, short) and lexical stress (+,-). For each of five categories 90 instances with possible /r/-realizations were extracted from a large speech database containing man-machine dialogues in an automatic train timetable inquiry system. The frequency of /r/-deletions in these 450 cases was investigated on the basis of variant selection by a CSR and human transcriptions of the same material. Loglinear analyses revealed that /r/-deletion was significantly more frequent after schwa than after full vowels, and that the effect of vowel length and lexical stress was not significant. This appeared from both the CSR data and the human transcriptions. Discrepancies between the two sets of results were observed, too. Possible explanations are discussed.

1. INTRODUCTION

In Dutch up to ten different realizations of postvocalic /r/ have been distinguished (Vieregge, Broeders, 1993; Van de Velde, 1996:130-131). Even more variants are presented elsewhere in this volume (Verstraete, Van de Velde, 2000). A perhaps somewhat unexpected variant of /r/ is its absence. In this contribution we will focus on the deletion of /r/ in postvocalic (coda) position before another consonant, since here /r/-deletion is most obvious in Dutch. We are, however, aware that /r/ is deleted in other positions as well, as was recently attested by Ernestus (2000:117-118).

Being tourists of our own language, as it were, we would first like to present some typical observations of /r/-deletions that we have come across in recent times.

1. Listening to man-machine dialogs in which train timetable information was

---

* This contribution is an extended version of Cucchiarini, van den Heuvel (1999). Agreement scores between the CSR data and the human transcriptions were added, and the Discussion section was extended with issues raised during the workshop and some other items that we consider relevant.

exchanged, we noticed that the common pronunciation of the Dutch capital *Amsterdam* was not the articulate form /Amst@rdAm/, but a reduced form /Ams@dAm/[1]

2. In a statistics syllabus we encountered the word *standaaddeviatie* which is a typo for *standaarddeviatie* (standard deviation).
3. Where the previous example is a typing error there are also cases where /r/-deletion entered the official orthography of Dutch. A typical example is the word *burgemeester* (mair), which has historically been derived from *burgermeester*.
4. In present-day written Dutch, hypercorrect forms such as *Spijkernisse* (placename; the correct spelling is: *Spijkenisse*) and *ofterwel* (or … either; the correct spelling is: *oftewel*) are, at times, observed.[2]
5. Some native speakers of Dutch have even reconstructed the word *slabber* as the stem of the diminutive *slabbetje* [slAb@tj@] which they think is spelled as *slabbertje* (bib). In fact, the correct spelling of the diminutive is *slabbetje* and its stem is *slab*.

The latter two examples clearly indicate that even native speakers easily associate a pronunciation without /r/ with a spelling in which the /r/ is present.

In a previous paper (Cucchiarini, Van den Heuvel, 1995), we accounted for the connected speech processes (CSPs) of postvocalic /r/-weakening and /r/-deletion in Dutch by referring to the linguistic context in which these phenomena are more likely to occur. On the basis of the data presented, it was concluded that a process of /r/-weakening is operative in Dutch connected speech, 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. An important factor to be acknowledged is the left vowel context (Cucchiarini, Van den Heuvel, 1995). The best way to test any hypothesis about these (and other) synchronic phenomena is to study them in real-life speech. For this reason in Cucchiarini, Van den Heuvel (1995) 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. Consequently, in a following paper (Cucchiarini, Van den Heuvel, 1998) we set out to obtain experimental evidence for the occurrence of postvocalic /r/-deletion in Dutch by checking how often /r/-deletion occurred in a large spontaneous speech corpus (214,102 words). To this end a

---

1 SAMPA phoneme notation will be used throughout this paper (Wells, 1997).

2 The latter example was reported by Piet van Reenen during the workshop.

continuous speech recognizer (CSR) in forced recognition mode was used. The CSR had the task to decide whether /r/-deletion was applied or not in a word.

The accuracy of forced recognition in selecting the correct variant was checked in an experiment (Kessens, Wester, Cucchiarini, Strik, 1998) in which the CSR's responses were compared with those of nine expert listeners who carried out the same task (i.e. deciding whether a segment is present or not). For /r/-deletion the CSR turned out to select the same pronunciation variant as the human transcribers did in 75.6% of the cases, while for the listeners this percentage varied between 74% and 93%. This indicated that there was a good correspondence between recognizer response and human judgements of /r/-deletion.

The results of forced recognition showed 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. Moreover, the results indicated that the frequency of occurrence of /r/-deletion was dependent on the left vowel context. First, we expected to find more instances of /r/-deletion when the left vowel is a schwa as opposed to any other kind of vowel. This was also confirmed by the analyses (Cucchiarini, Van den Heuvel, 1998). Second, we thought the frequency of /r/-deletion may vary depending on whether the preceding vowel is short or long, for the following reason. If an /r/ is deleted after a short vowel in polysyllabic words, then some restructuring has to take place because a short vowel is not allowed in syllable-final position in Dutch (Booij, 1995:5). This requirement might have an inhibiting effect on the application of /r/-deletion, with the consequence that /r/-deletion is more frequent after long vowels than after short vowels, at least in polysyllabic words However, the analyses showed that vowel length had no significant effect on /r/-deletion frequency. A possible explanation for this finding could be that in the experiment reported on in Cucchiarini, Van den Heuvel (1998) the effect of vowel length was confounded with that of stress. If we consider that schwa is “unstressable” in Dutch (Booij, 1995:20), then the findings presented in Cucchiarini, Van den Heuvel (1998) suggest that the phenomenon of /r/-deletion might be related to stress: it is more frequent after a vowel that is never stressed like schwa than after any other vowel that can potentially bear stress. Since in the experiment in Cucchiarini, Van den Heuvel (1998) no distinction was made between stressed and unstressed short and long vowels, the relation between stress and /r/-deletion could not be investigated.

In order to study these factors separately, we decided to carry out another experiment that was directed at studying the phenomenon of /r/-deletion under various conditions of left vowel context including its stress property. Furthermore, although the performance of the CSR had

been shown to be comparable with that of expert listeners, we thought evidence in terms of detailed phonemic transcriptions from human transcribers was needed for a good understanding of the relationship between /r/-deletion and left vowel context. Therefore, the speech material used as input to the CSR was also transcribed by thirteen transcribers.
To summarize, the aim of the experiment reported in this paper is to determine whether /r/-deletion is dependent on the length and the degree of stress of the preceding vowel, or whether the distinction between schwa and full vowel is the only determinant of this phenomenon. To establish this two different types of evidence were gathered: evidence from the performance of a CSR and evidence from human transcribers. Both types of data will be presented and analyzed in the rest of this paper.

2. METHOD

2.1. Speech Material and Design

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 stems from man-machine interactions in an automatic train time-table inquiry system (Strik, Russel, Van den Heuvel, Cucchiarini, Boves, 1997). The speech style of this material can be characterized as semi-spontaneous, since each caller has a general idea of the intended query to the machine without having prepared the exact wordings. The waveform format of the speech files is A-law, sampled at 8 kHZ. The VIOS1 training database was used to train the CSR (see further section 2.3). The test utterances were selected so as to obtain a full factorial design for the relevant effects (vowel quality, length and stress). Accordingly, five classes for the left vowel of /r/ were distinguished: 1. schwa (which is always unstressed); 2. short vowel, stressed; 3. short vowel, unstressed; 4. long vowel, stressed; 5. long vowel, unstressed. Note that before /r/ only five vowels in Dutch are short: /I, Y, E, A, O/. For each category 90 samples (target words) were chosen, giving a total of 450 samples. This rendered the task feasible for the human transcribers, while still permitting sufficient samples for statistical analysis of the data.
The 450 target words were taken from utterances that were not in the training set. Furthermore, care was taken that the target words were not monosyllabic. In monosyllabic words with a short full vowel, syllable reorganization is not required because the short vowel would not be syllable-final anyway. The 450 target words selected stemmed from 425 unique utterances from 385 different dialogues.

2.2. Automatic Variant Selection

On the basis of the /r/-deletion rule as specified in Cucchiarini, Van den Heuvel (1995, 1998), variants with /r/ and variants without /r/ were generated for all relevant words in the test set in which the relevant contexts were met, by means of a Perl implementation of the rule in question. For each of these words the pronunciation variants with and without /r/ were included in the lexicon that was used for forced recognition. In forced recognition mode the CSR is not used to 'recognize' which word was spoken, but it 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, because this is already known, but which pronunciation form of that word was realized, as in this study.

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 used 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 Strik et al. (1997). The phoneme models were trained on the training material of the VIOS1 database containing 25,104 utterances stemming from 3,530 different dialogues.

2.3 Human Transcriptions

Thirteen transcribers were asked to transcribe the above-mentioned 450 realizations of postvocalic /r/. The transcribers were Language and Speech Pathology students at the University of Nijmegen. They were all female, and had attended the same transcription course including 32 hours contact time. The transcription system used in this course is that of the International Phonetic Association (IPA).

The transcribers worked in small groups of two or three people (five duos and one trio) and based their transcriptions on auditory analysis of the full utterances without any kind of visual support. Each group produced a consensus transcription of one sixth of the material (75 realizations). The task was to transcribe only the target /r/. The transcribers were not just instructed to determine whether they heard an /r/ or not, but they had to use the full range of

IPA symbols and diacritics at their disposal to describe their auditory impression. For the analyses presented below we took a conservative stance and interpreted only full deletions of /r/ (or zero-transcriptions) as factual /r/-deletions.

3. RESULTS

3.1 General Findings

The transcriptions of five /r/-realizations turned out to be missing, so that we had to remove these cases from our analyses. Table 1 presents the frequencies and percentages of /r/-deletions in all categories both for the human listeners and for forced recognition by the CSR.

		CSR		Human	
	Possible	Applied	perc.	applied	Perc.
After schwa	89	50	56%	30	34%
After short vowel	179	40	22%	24	13%
+ stress	*90*	*19*	*21%*	*12*	*13%*
- stress	*89*	*21*	*24%*	*12*	*13%*
after long vowel	177	47	27%	13	7%
+ stress	*90*	*21*	*23%*	*5*	*6%*
- stress	*87*	*26*	*30%*	*8*	*9%*

Table 1. Number of possible applications, number of applications and percentage of application of /r/-deletion in the various contexts, for both the machine and the human responses.

For the CSR responses Table I shows that the type of vowel preceding /r/ has an effect on the amount of deletion: /r/-deletion is clearly more likely to occur after schwa than after any other type of vowel, as already observed in Cucchiarini, Van den Heuvel (1998). However, the effect of the factor stress is not very clear: for short vowels there is hardly any difference, while for long vowels there seems to be more deletion after unstressed vowels.

Table 1 further shows that the human listeners scored fewer cases of /r/-deletion, overall, than the CSR did. It can also be seen that the scores of the human transcribers show the same behavior as those of the CSR. After schwa /r/-deletion occurs much more frequently than after full vowels, whereas the factors length and stress within the full vowels do not seem to influence /r/-deletion markedly. We find a tendency for long vowels to have somewhat more /r/-deletions after unstressed vowels than after stressed vowels, which was also observed for the CSR.

The above-mentioned impressions for the /r/-deletions in the human and the CSR data sets have to be substantiated by a series of more thorough statistical analyses. The next subsections report on these analyses.

3.2 Automatic Variant Selection

Loglinear analyses were performed since they are typically suited to deal with frequency data in more complex factorial designs (Rietveld, Van Hout, 1993). The first impression to be tested for our data is the hypothesis that factors 'vowel length' and 'stress' do not have a significant impact on /r/ deletion after full vowels. To test the effects of the two factors a hierarchical loglinear analysis was carried out on a subset of the data. The contingency table used as input to loglinear analysis is given in Table 2.

	short vowel		long vowel		Total
	+ stress	- stress	+ stress	-stress	
- r	19	21	21	26	87
+ r	71	68	69	61	269
Total	90	89	90	87	356

Table 2. Contingency table for full vowels with frequencies of /r/-deletion (-r) and /r/-retention (+r) as affected by vowel length and stress. Scores for CSR in forced recognition mode.

In carrying out loglinear analysis we started with a saturated model, that is one containing all possible effects: '/r/-deletion', 'vowel length', 'stress' and the four possible interactions. The results of this analysis show that none of the interaction effects is significant; a significant effect was found only for the '/r/-deletion' factor ($z = -9{,}11$; $p<0.001$). In other words, the frequency of /r/-deletion does not seem to be dependent on the length of and/or the amount of stress on the preceding vowel.

In the light of these findings we decided to pool the data pertaining to the groups of stressed and unstressed short and long vowels so as to form a new category 'full vowel' which could then be compared with schwa. Again a hierarchical loglinear analysis with a saturated model was carried out, this time with the factors '/r/-deletion', 'vowel quality' (with the two levels 'full vowel' and 'schwa'), and the interaction between the two factors.

Table 3 shows the data submitted to loglinear analysis. The results of this analysis show that only a saturated model is appropriate, since both effects and their interaction are significant: $z=5.58$ ($p<0.001$) for the interaction effect, $z=-3.58$ ($p<0.001$) for /r/-deletion and $z=-10.06$ ($p<0.001$) for vowel quality. In other words, the difference in degree of /r/-deletion observed between schwa and the other vowels appears to be statistically significant.

	full vowel	Schwa	Total
- r	87	50	137
+ r	269	39	308
Total	356	89	445

Table 3. Contingency table for full vowels and schwa with frequencies of /r/-deletion (-r) and /r/-retention (+r) as affected by vowel quality. Scores for CSR in forced recognition mode.

3.3. Human Transcriptions

To substantiate our impressions of the human transcriptions we performed two hierarchical loglinear analyses with the same designs as those outlined in section 3.2. First, the interaction effect of '/r/-deletion' with factors 'vowel length' and 'stress' for the full vowels was tested with the data in Table 4.

	short vowel		long vowel		Total
	+ stress	- stress	+ stress	-stress	
- r	12	12	5	8	37
+ r	78	77	85	79	319
Total	90	89	90	87	356

Table 4 Contingency table for full vowels with frequencies of /r/-deletion (-r) and /r/-retention (+r) as affected by vowel length and stress. Scores found for human transcribers.

Similar to the CSR scores, only the factor '/r/-deletion' *per se* was significant ($z = -12.11$; $p < 0.001$), whereas the interactions with factors 'vowel length' and 'stress' were not. For this reason, we pooled the data for the full vowels and compared them with the /r/-deletions after schwa, as displayed in Table 5.

	full vowel	Schwa	Total
- r	37	30	67
+ r	319	59	378
Total	356	89	445

Table 5. Contingency table for full vowels and schwa with frequencies of /r/-deletion (-r) and /r/-retention (+r) as affected by vowel quality. Scores found for human transcribers.

The corresponding hierarchical loglinear analysis revealed that only the saturated model fitted the data appropriately. The interaction between '/r/-deletion' and 'vowel quality' proved significant at $p=0.001$ ($z=5.23$). This means that the difference in degree of /r/-deletion observed between schwa and the other vowels is statistically significant, which replicates the outcome for the CSR transcriptions.

3.4 Pairwise comparison of transcriptions

Although both human and CSR data show the same pattern in the distributions of the /r/-deletions, it might be interesting to know to what extent CSR and humans choose the same alternative for individual word tokens. This kind of analysis is not so much required to understand the phenomenon of /r/-deletion, but it may be useful to get more insight into the differences in performance between humans and machine.

From the data presented above we can infer that the number of cases in which CSR and humans make the same choice cannot be very high as the CSR finds many more deletions than the human listeners do. To get a clearer idea we computed the degree of agreement between the CSR and the human transcribers. Again, only full deletions and 0-symbols were counted as /r/-deletions in the human transcriptions.

The degree of agreement is expressed both in percentage agreement and in Cohen's κ. We chose percentage agreement because it is a widely used measure. However, for data of this kind in which the percentage of occurrence of a given phenomenon is very low (there are relatively few /r/-deletions for both CSR and humans) percentage agreement may be inflated by chance agreement (Suen, Ary, 1989) and it is therefore more appropriate to use a metric which corrects for chance agreement such as Cohen's κ (Rietveld, Van Hout, 1993). For the full set the overall percentage agreement was 76.6% and Cohen's κ was 0.36 (z=8.43). Table 6 lists the agreement values for the five categories.

	+stress	-stress
Short vowel	0.34 (80.7%)	0.45 (83.1%)
Long vowel	0.24 (80.0%)	0.32 (77.3%)
Schwa	-	0.26 (61.4%)

Table 6. Agreement in terms of Cohen's κ between CSR data and human transcriptions for the five categories. The corresponding percentage agreements are added in brackets.

Table 6 shows clearly that, although percentage agreement between the two sets of transcriptions is rather high, agreement in terms of κ is low. As explained above, we did not expect a high agreement coefficient because we already knew that the CSR opted for many more /r/-deletions than the humans did. Different factors may have led to these results.

First, the task carried out by the transcribers differed essentially from the CSR's task. The transcribers could choose from among many possibilities, while the CSR's choice was dichotomous. In other words, the CSR was forced to choose between a full /r/ and a deleted /r/, whereas the transcribers could indicate various degrees of weakening by using diacritics or other phonetic symbols like schwa. In this context, we refer to Kessens et al. (1998) where the human transcribers only had a dichotomous choice in the sense that they had to judge whether a specific phone was present or absent in a speech signal. Indeed much higher agreement values were reported for that experiment ($\kappa = .52$).

For the present experiment, the transcribers indeed transcribed a schwa instead of an /r/ in many cases. One might argue that these occurrences should be counted as instances of /r/-deletion, especially if schwa is the left vowel context. In fact, by considering schwa transcriptions as /r/-deletions the responses from the CSR and the human transcribers look quite alike in terms of absolute frequencies. However, for the present analyses we decided to adopt a more conservative approach, and limited ourselves to factual /r/-deletion cases. In the near future we intend to analyze the transcription data in more detail by studying other mappings in order to see how they differ from the CSR's data.

A second possible explanation for the lower percentages of deletion in the human data is that human transcribers are likely to be influenced by their knowledge of the orthographic representation of words (Cucchiarini, 1993). Since in this experiment the transcribers listened to whole utterances, they knew which words the speaker was uttering and this might have induced them to actually 'hear' an /r/ when in fact it was not there. However, we should not forget that transcribing connected speech processes in categorical terms may be very difficult in certain cases, as has been pointed out by Booij (1995:126). Typically, certain phenomena, like the one discussed in this paper, may exhibit a gradual nature, so that in certain cases it can be very difficult to determine whether /r/ is extremely weakened or completely deleted. Nonetheless, it would be interesting to have the same material transcribed by phoneticians who are unfamiliar with Dutch. Thus, the influence of orthographic knowledge can be excluded.

Another possibility is that the discrepancy between the CSR data and the human transcriptions in 'perceiving' /r/-deletion is related to the very nature of the HMMs employed in the CSR. These models are essentially static, in the sense that they do not take much account of neighboring sounds. With respect to human perception, on the other hand, we know that the way one sound is perceived very much depends on the identity of the adjacent sounds and the transitions between the sounds. If the cues for a given phone are contained in adjacent sounds, the phone in question would appear to be present to human listeners, but would be absent for the CSR which does not make use of such cues.

Finally, a possible explanation is that the CSR and the transcribers have different durational thresholds for detecting an /r/, in the sense that phones with a duration that falls under a certain threshold are less likely to be detected. This sounds plausible if we consider the topology of the HMMs used in this study. Our HMMs consist of three states and this means that phones that last less than 30 ms are less likely to be detected (feature extraction is done every 10 ms).

To get a better understanding of the differences between the CSR's choices and those of the transcribers against the background that the CSR chooses more deletions, it might be interesting to know to what extent the deletions that the humans did find were also found by the CSR. Of course this is something that affects the value of κ, but this information cannot be readily extracted from a κ value. For this reason we computed the percentage of deletions found by the CSR for the cases in which the transcribers chose /r/-deletions. Table 7 presents these results.

	+stress	-stress
Short vowel	58% (=7/12)	75% (=9/12)
Long vowel	80% (=4/5)	87.5% (=7/8)
Schwa	-	77% (=23/30)

Table 6. Percentage of deletions found by the CSR for the cases in which the transcribers opted for /r/-deletion. The corresponding raw scores are given in brackets.

As can be seen, in the majority of cases in which the transcribers found a deletion, the CSR also found a deletion. This result is reassuring since it indicates that the differences between the CSR and the transcribers are not random, but can probably be attributed to one or two parameters along which the CSR and the transcribers differ, for instance the durational thresholds or the use of cues from neighboring sounds. In any case further insight into these differences may be obtained by experimenting with different mappings between the CSR's response and those of the transcribers.

4. DISCUSSION AND CONCLUSIONS

In this paper we have examined the process of postvocalic /r/-deletion in Dutch from two different perspectives: that of pronunciation variants selected by a CSR and that of phonetic transcriptions made by human transcribers. With respect to the factors influencing the /r/-deletion process both sets of data produce the same results. In particular, the quality and the characteristics of the left vowel context appear to be relevant only to a certain extent for /r/-deletion: if the preceding vowel is a schwa, /r/-deletion is favored, whereas if this vowel is a full vowel, /r/-deletion is less likely. For the rest no other distinction within the 'full vowel' category seems to be relevant, neither in length nor in degree of stress. This pattern of results also seems to be rather stable, as it emerged from both types of transcriptions.

These findings would seem to suggest that the stress distinction that is relevant to the process of /r/-deletion is not a gradual one, but a dichotomous one in terms of stressable *vs* unstressable, with "unstressability" favoring /r/-deletion. Since distinctions are fainter in an unstressed environment (Boersma, 1998:210), such ultimately unstressed syllables, like those containing schwa, can more easily tolerate further reduction, like that of postvocalic /r/, than any other type of syllable. For the full vowels the stress distinction appears to be less clear-cut. Although unstressed vowels will always be realised as weak, the prominence of stressed vowels may depend on the accent pattern of the utterance.

As for the factor 'vowel length', the CSR data tend towards a higher proportion of /r/-deletions after long vowels than after short vowels, whereas the opposite tendency is seen for the human transcriptions. Perhaps this tendency becomes significant if a larger sample is used, but we cannot properly test this, since a power analysis on our statistics is not possible: such a procedure does not presently exist for loglinear analyses.

A factor which was recently reported to have a possibly significant influence is 'vowel height' (Ernestus, 2000:117-118). The hypothesis here is that low vowels favor the deletion of ensuing /r/. Our data set was not designed for this analysis, but we consider it a relevant suggestion for further research.

Although the two sets of data are in concordance as to the impact of the factors under investigation, it is clear that they differ with respect to the percentage of occurrence of /r/-deletion. More precisely, the CSR data reveal higher percentages of /r/-deletion than the human transcriptions. Possible explanations for this finding have been discussed and will be tested in future investigations.

To summarize, the data presented in this paper reveal some differences in percentages of /r/-deletion between the CSR data and the human transcriptions. These differences, however, do not interfere with the main aim of this investigation. Our main objective was to determine which properties of the left vowel context favor postvocalic /r/-deletion. In this respect both types of data examined yield the same results: /r/-deletion turns out to be much more common after schwa than after any full vowel, irrespective of the length and stress characteristics of that vowel.

ACKNOWLEDGEMENTS

We would like to thank Mirjam Wester and Judith Kessens for preparing the tapes for the human transcribers, and for carrying out forced recognition and making the results available to us. We would also like to thank Judith for computing the κ scores reported in section 3.4. We thank Felix Vieregge for kindly integrating our transcription material in his course curriculum.

REFERENCES

Boersma, P. (1998). *Functional phonology. Formalizing the interactions between articulatory and perceptual drives.* PhD Thesis. University of Amsterdam.

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

Cucchiarini, C. (1993). *Phonetic transcription: a methodological and empirical study.* PhD Thesis, University of Nijmegen.

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

Cucchiarini, C., Van den Heuvel, H. (1998). /r/-deletion in Standard Dutch. How phonological research can profit from ASR technology. *Proc. ICSLP'98,* 5: 1759-1762.

Cucchiarini, C., Van den Heuvel, H. (1999). Postvocalic /r/-deletion in Dutch: more experimental evidence. *Proc. ICPhS*, 1999, 3: 1673-1676.

Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch.* PhD Thesis. University of Amsterdam.

Kessens, J., Wester, M., Cucchiarini, C., Strik, H. (1998). The selection of pronunciation variants in spontaneous speech: Comparing the performance of man and machine. *Proc. ICSLP'98*, 6: 2715-2718.

Rietveld, T., Van Hout, R. (1993). *Statistical techniques for the study of language and* *language behaviour*. Berlin, New York: Mouton de Gruyter.

Suen, H.K, Ary, D. (1989). *Analyzing quantitative behavioral observation data*. Hillsdale (NJ): Lawrence Erlbaum.

Strik, H., Russel, A., Van den Heuvel, H., Cucchiarini, C., Boves, L. (1997). A spoken dialogue system for the Dutch public transport information service. *International Journal of Speech Technology*, 2:121-131.

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

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

Verstraete, B., Van de Velde, H. (2000). Socio-geographical variation of /r/ in standard Dutch. This volume.

Vieregge, W.H. and Broeders, T. (1993). Intra- en interspeaker variation of /r/ in Dutch, *Proc. Eurospeech'93*, Berlin, 267-270.

Wells, J. (1997). Standards, Assessments, and Methods: phonetic alphabets. University College London. See also: http://www.phon.ucl.ac.uk/home/sampa/home.htm