HALF A WORD IS ENOUGH FOR LISTENERS, BUT PROBLEMATIC FOR ASR

H. Strik, A. Elffers, D. Bavcar, C. Cucchiarini

Centre for Language and Speech Technology (CLST)
Radboud University Nijmegen, the Netherlands

ABSTRACT

The present study investigates whether there are word sequences that exhibit considerable deviation in pronunciation, such that they might require special treatment in speech technology, so called multiword expressions (MWEs). The results show that these sequences exist, that they are frequent and that they are often extremely reduced. In order to be studied, MWEs have to be identified in the first place. We investigate how such sequences can be automatically detected in a corpus of spontaneous speech. Measures that are known to be related to predictability and phonetic reduction are employed for this purpose. Our findings indicate that these measures yield different results and that a combination of criteria would probably be most effective.

1. INTRODUCTION

Over the last decades the performance of automatic speech recognition (ASR) and spoken dialogue systems has improved considerably, but handling spontaneous speech and dialogues still turns out to be problematic. One of the reasons is that spontaneous speech contains a lot of variation which cannot be dealt with adequately by current models. Although pronunciation variation modeling for ASR has received substantial attention [20, 22], many things remain unclear and some issues have barely been touched upon. Multiword expressions (MWEs) are a case in point. They have already been studied quite extensively in NLP [see, e.g., 11, 13, 14, 18]. However, the pronunciation of MWEs has received very little attention.

In our own research we have studied pronunciation variation in word sequences that may qualify as MWEs [3, 21]. Since there is no generally accepted definition of the notion of MWE in spoken language, we based our investigations on what we consider a reasonable operational definition of this concept: MWEs are sequences of words that are characterized by unpredictable pronunciation.

In [3, 21], frequent word sequences were first extracted and then further analyzed in that part of the 'spoken Dutch corpus' (Corpus Gesproken Nederlands: CGN [15, 16, 25]) that comes with manually verified broad phonetic transcriptions. The results show that many of these word sequences exhibit uncommon pronunciation patterns that are not found in other contexts and that usually the words in such MWEs are (much) more reduced than the same words in other contexts.

MWEs are thus characterized by a considerable amount of reduction [3, 21], which can be problematic for ASR, if not handled properly. However, it remains to be seen whether deviant pronunciation patterns are a prerogative of such highly frequent stock phrases or whether they are also encountered in less frequent sequences that are not readily recognized as being stock phrases. In this connection it may be worthwhile to refer to research by [1, 8, 9], which has shown that predictable words are more likely to be reduced. One can imagine that there may be word sequences that are not readily categorized as stock phrases, but that occur frequently enough as to exhibit high predictability and therefore considerable deviation in pronunciation.

Another relevant finding in this respect is that fixed expressions occur frequently in spontaneous speech. In [3] we found that 21% of the source corpus investigated consisted of fixed expressions. As cognitive load increases, speakers are more likely to use prefabricated expressions [12, 17]. In commentaries of sports games such expressions can cover up to 48% of the whole speech material [17].

The question that arises at this point is not only whether word sequences with deviant pronunciation exist, but also how they can be detected automatically, because in the end this is the only way that data from large corpora can be handled to the benefit of research into speech science and speech technology.. In other words, which criteria can be applied to spot such sequences in a corpus of spontaneous speech? In this paper we address these issues on the basis of a study on the CGN spontaneous speech in which a number of possible indicators of deviant pronunciation are investigated to determine which of them are most promising for selecting potential MWEs.

2. MATERIAL

The corpus used for the current study is CGN, a database containing about 9 million words of contemporary Dutch as spoken in the Netherlands and Flanders [15, 16, 25]. All recordings are orthographically transcribed, lemmatized and enriched with part-of-speech (POS) information.

For about 10% of the corpus, more detailed annotations are available, such as manually checked broad phonetic transcriptions, word alignments, and syntactic and prosodic annotations. For the phonetic transcriptions a computer phonetic alphabet was used [25] that is a slightly modified version of SAMPA [26]. This sub-corpus of 900,000 words, called the core corpus, was composed in such a way that it faithfully reflects the design of the full corpus. In this paper we report results for spontaneous dialogues (component A) of the core corpus, 100.989 words in total.

3. CASE STUDY: "OP EEN GEGEVEN MOMENT"

In our previous research we identified word sequences that may qualify as MWEs [3, 21] because they display deviant behavior in pronunciation. In order to come to a better understanding of this behavior we studied some of these MWEs in more detail. In this section we will present the results for the MWE "op een gegeven moment".

In component A of the core corpus 22 occurrences of "op een gegeven moment" were found. They are presented in table 1. It can be observed that among the 22 transcriptions there are 20 different ones, only two transcriptions occur twice (see *'s in table 1).

The canonical transcription of "op een gegeven moment" is /Op en G@gev@ momEnt/ (16 phonemes and 7 syllables). All 22 transcriptions are shorter than the canonical form, indicating reduction, and in many cases the reduction is substantial. In order to get an idea of the amount of reduction, all 22 transcriptions in table 1 were compared to the canonical transcription by means of a DP-program which calculates the number of substitutions, deletions, insertions, and percentage disagreement [4]. There are no insertions, many deletions and some substitutions (usually vowel reduction). On average there are 6.4 deletions and 1.2 substitutions, which means that almost half of the phonemes are not pronounced in the canonical way. There are also some cases in which the number of changes is 9, 10 or 11, compared to the 16 phonemes of the canonical pronunciation. The average percentage disagreement is 47.7%, ranging from 18.8% to 68.8%. The number of syllables, seven in the canonical pronunciation, is sometimes reduced to four, three, or even two.

Table 1. The 22 realisations of "op een gegeven moment".

Realisation	Sub	Del	Ins	%Dis
Op @n xe m@nt	1	6	0	43.8
Op @ xe m@nd	2	7	0	56.3
Op @ x@f mEnt	1	6	0	43.8
Op @N Gev@ momEnt	1	2	0	18.8
Ob @ xev@ mEnt	1	5	0	37.5
p @ Ge md	1	10	0	68.8
Op @ Ge mt	0	9	0	56.3
Op @ xe mnt	0	8	0	50.0
Op @ Ge m@nt	1	7	0	50.0
Ob @ Ge m@t	2	8	0	62.5
Op @ Ge @nt	1	8	0	56.3
@b @ Gev mEnt	2	6	0	50.0
Op @ xev@ m@nt *	1	5	0	37.5
Op @ xev@ m@nt *	1	5	0	37.5
Op @ xe m@n	1	8	0	56.3
Ob @N xev@ mEnd	3	4	0	43.8
Ob @ Ge m@n	2	8	0	62.5
Ob @ Ge m@nt	2	7	0	56.3
Ob @ Gev m@nt	2	6	0	50.0
ub @ Gev mEnt	2	6	0	50.0
Op @ Gev@ mEnt *	0	5	0	31.3
Op @ Gev@ mEnt *	0	5	0	31.3
Average	1.2	6.4	0.0	47.7

In table 1 the transcriptions of 22 realizations are compared to the canonical transcription. It can be observed that the pronunciation of the words in these realizations differs from the canonical one. Besides comparing the transcription of realizations to canonical transcriptions, we also compared words in MWE contexts to the same words in other contexts. Values for the following measures are presented here:

- FRQ = frequency of occurrence
- LEN = length (number of phonemes)
- DUR = duration (msec.)
- ART = articulation rate (LEN/DUR)

In table 2 some values are presented for these measures: average values in MWE context, in all other contexts, and the differences between the two. Note that for the word "gegeven" there is only one occurrence in other context. This is observed also for words in other MWEs, i.e. that they occur almost solely in the context of a MWE and rarely or not at all in other contexts.

It can be observed that for all words the average values of length and duration in the MWE context are smaller than the corresponding ones in other contexts. Conversely, the articulation rate values are higher in the MWE context than in the other contexts (because, on average, the decrease in duration is larger than the decrease in length). These results clearly indicate that in the MWE context all words are (much) more reduced than in other contexts, which is a plausible finding. In some cases the degree of reduction is very high: in 10 of the 12 cases in table 2 the percentage difference is more than 24% and the maximum is a reduction by 52% (duration of word 'een'). Note that these are average values, which means that there are cases in which the amount of reduction is even higher.

Table 2. Measurements for the words in the MWE "op een gegeven moment"; for LEN, DUR and ART average values are presented in MWE context, in all other contexts, and the differences between the two.

		Op	een	gegeven	moment
FRQ	MWE	22	22	22	22
	Other	325	1470	1	24
LEN	MWE	1.95	1.14	2.82	3.68
	Other	1.97	1.82	4.00	5.63
	Diff	-0.02	-0.69	-1.18	-1.94
	%Diff	-1%	-38%	-30%	-35%
DUR	MWE	0.09	0.05	0.19	0.17
	Other	0.12	0.11	0.29	0.34
	Diff	-0.03	-0.06	-0.10	-0.16
	%Diff	-24%	-52%	-35%	-49%
ART	MWE	20.6	20.8	14.8	21.2
	Other	15.8	16.1	13.7	16.7
	Diff	4.8	4.6	1.1	4.5
	%Diff	30%	29%	8%	27%

It is obvious that the pronunciation of words in MWEs differs substantially from the pronunciation of the same words in other contexts. As is well known, pronunciation variation is likely to be problematic for current ASR systems, leading to higher error rates. This suggests that proper handling of MWEs is required to enhance the performance of ASR systems.

4. EXTRACTION OF MWES

In order to find out how MWEs can best be handled in ASR systems, it is necessary to study how MWEs behave and how they can be represented in ASR systems. This sort of study, however, requires that MWEs be first identified and extracted from speech corpora. Since it is not immediately clear how this should be done, we decided to study whether there are measures that can be used to identify potential MWEs. Some results of this part of the study are presented here.

We wondered whether it would be possible to detect MWEs automatically by resorting to some measure of reduction that can be easily calculated automatically. Since various studies have revealed that reduction is related to frequency [5, 24], the first obvious measure to consider would be frequency of occurrence. Second, although we observe that words in multiword contexts exhibit a considerable amount of vowel reduction, i.e. substitutions of full vowels in the canonical transcriptions by schwas in the actual pronunciations, it clearly appeared that the amount of reduction in pronunciation was mainly caused by the fact that many segments in the canonical representation turned out to be deleted in the actual pronunciation. On the basis of these findings one would think that a length measure that expresses the relation between the number of segments in the canonical representation and the number of segments in the actual pronunciation should be a good indicator of reduction and deviant pronunciation. For this purpose we devised the following two length measures:

- ALD (absolute length difference) =
  # segments in canonical representation –
  # segments in actual pronunciation
- RLD (relative length difference) = 100% *
  ALD / (# segments in canonical representation).

The ALD measure is likely to select longer N-grams, while for the RLD this is probably not the case. Since these measures will have to be compared for the various N-grams, one could argue that a relative measure such as RLD is to be preferred for comparability reasons. On the other hand, a measure like RLD will not be able to reveal cases of extreme reduction. For example, RLD will make no distinction between the following two cases A and B:

A: RLD = 100%*(4–2) /4 = 50%,
ALD = 4–2 = 2

B: RLD = 100%*(20–10) /20 = 50%
ALD = 20–10 = 10

However, distinguishing between these two cases may be extremely relevant for purposes such as pronunciation variation modeling and automatic transcription, for the simple reason that cases in which only two symbols are deleted can probably be accounted for in the form of rewrite rules applied to individual words when generating a multi-pronunciation lexicon. However, in the case in which many segments are deleted and especially if these segments are deleted in clusters, it is unlikely that the deviant pronunciation pattern can be accounted for by rewrite rules. These are the cases that will probably require that the N-grams in question be treated as lexical entries in the pronunciation lexicons used in ASR and automatic phonetic transcription, with their own specific pronunciation variants.

We also look at other measures for selecting MWEs. For instance, it has long been known that faster speech generally leads to a greater amount of reduction [7]. This is corroborated by the results we have obtained for MWEs, (table 2). Therefore, it would seem plausible to use some measure of speaking rate as an indicator of reduction. In our experiments we investigated the potential role of articulation rate (in isolation), but this did not turn out to be a good measure for selecting MWEs. For this reason, these results are not presented here.

4.1. FRQ: Frequency of occurrence

We now look at the N-grams selected on the basis of the frequency criterion. The ten most frequent N-grams in table 3 all consist of 2 words. It is obvious that, on average, shorter N-grams will occur more frequently than longer ones. Therefore, it is questionable whether frequency alone is an appropriate measure for selecting MWEs. Frequency remains important, of course, since better modeling frequent events is more likely to improve ASR performance than improving the modeling of less frequent events. Frequency should therefore be used in combination with other selection measures.

Table 3. The ten N-grams with the highest FRQ values.

	Orthography	FRQ	ALD	RLD	ART
1	ja ja	442	-0.03	-0.80	9.2
2	dat is	367	0.40	7.90	15.2
3	ja maar	271	0.61	12.3	12.0
4	da's	266	-2.62	-87.2	16.8
5	en dan	244	0.46	9.10	17.2
6	ja dat	226	0.42	8.40	12.5
7	't is	223	0.92	23.0	15.0
8	of zo	212	0.03	0.70	12.3
9	als je	209	1.97	39.3	12.8
10	oh ja	207	-0.03	-1.13	7.90

4.2. ALD: absolute length difference

In table 4 the ten N-grams with the highest ALD values are presented, in descending ALD order. The maximum ALD is seven, indicating that for the three occurrences of this N-gram the actual pronunciation contains, on average, seven segments less than the canonical representation. Table 4 reveals that in general there is a large amount of reduction and that it often concerns N-grams containing the sequence "gegeven moment". The N-grams with rank order 1, 2, 5 and 10 are all different subsets of the same sequence ("je op een gegeven moment ook") that occurs three times. Furthermore, it appears that there is no overlap between the N-grams listed in table 4 and those in table 3.

Table 4. The ten N-grams with the highest ALD values.

	Orthography	FRQ	ALD	RLD	ART
1	op een gegeven moment ook	3	7.00	38.9	20.4
2	een gegeven moment ook	3	6.67	41.7	20.5
3	op een gegeven moment	22	6.41	40.1	18.7
4	een gegeven moment	22	6.36	45.5	18.3
5	gegeven moment ook	3	6.00	42.9	19.8
6	natuurlijk helemaal	3	6.00	40.0	20.9
7	dan op een gegeven moment	3	5.67	29.8	20.4
8	is in ieder geval	3	5.67	43.6	18.4
9	gegeven moment	28	5.39	44.9	18.6
10	je op een gegeven	3	5.33	44.4	19.3

4.3. RLD: relative length difference

Table 5 shows the ten N-grams with the highest RLD values in descending RLD order. The highest RLD value is almost 50%, meaning that, on average, for the 27 occurrences of the N-gram "een gegeven" the actual pronunciation contains about half the number of segments present in the canonical transcription. In addition, we observe no overlap between the N-grams listed in table 5 and those in table 3, whereas there is some overlap between tables 5 and 4.

Table 5. The ten N-grams with the highest RLD values.

	Orthography	FRQ	ALD	RLD	ART
1	een gegeven	27	3.96	49.5	16.2
2	een gegeven moment	22	6.36	45.5	18.3
3	gegeven moment	28	5.39	44.9	18.6
4	je op een gegeven	3	5.33	44.4	19.3
5	hè als	9	2.22	44.4	9.07
6	als je als	6	3.50	43.8	10.7
7	is in ieder geval	3	5.67	43.6	18.4
8	ze natuurlijk	3	4.33	43.3	20.3
9	gegeven moment ook	3	6.00	42.9	19.8
10	een gegeven moment ook	3	6.67	41.7	20.5

Among the N-grams with a high RLD value, there are many N-grams containing the word "gegeven", as was the case for those selected on the basis of high ALD values. Since the N-grams containing the word "gegeven" appear to have such a deviant pronunciation pattern, we decided to study them in more detail.

5. GENERAL DISCUSSION

The results presented above and those of previous research make it clear that there are word sequences that qualify as MWEs because they exhibit deviant pronunciation behavior. In particular, this deviant pronunciation is generally characterized by otherwise unusual forms of reduction. Up till now, this type of variation has received little attention and is generally not represented in current models.

Modeling pronunciation variation for ASR has received substantial attention [20, 22]. Although different methods have been used, the majority of them uses a similar strategy which can shortly be described as follows. Pronunciation variants with different transcriptions are added to the lexicon to model differences at the symbolic level, while the acoustic models mainly model the acoustic differences between the various occurrences of the symbols (e.g. by incorporating models for triphones or other N-phones, and by using Gaussian mixtures).

The amount of pronunciation variation observed for MWEs is large. Not only does a lot of reduction occur, but a lot of different reduced variants can be observed ranging from citation forms to extremely reduced forms (e.g., 20 variants occur for the MWE "op een gegeven moment"). Adding all the variants of the individual words of the MWEs to the lexicon is not likely to work in all cases. In [10] we saw that this can be counterproductive. Adding complete MWEs and their pronunciation variants to the lexicon produced better results in [10]. Others also found that adding multiwords to the lexicon and treating these as words with their own specific pronunciation variants can improve ASR performance [e.g., 2, 6, 19]. Since in general such studies were mainly aimed at reducing word error rate in ASR, usually by simply adding multiwords and their pronunciation variants to the lexicon, the behavior of the MWEs was not studied in detail. The criterion used most often to select the multiwords was frequency of occurrence. However, in these studies the number of variants added and the amount of reduction modeled was limited.

To summarize, there are indications that MWEs occur frequently, that they are substantially reduced, and that they exhibit many different variants. Future research seems to be required to be able to shed more light on how MWEs should be handled. However, to make such research possible it is also necessary to automatically find out which word sequences are potential MWEs. To this end different measures were investigated in this paper. The results show that when these measures are applied to identify N-grams with reduced pronunciation, different N-grams are selected. Two of these measures, ALD and RLD, show some degree of overlap and turn out to select similar N-grams, which is not surprising given that both measures express a difference in length between the canonical representation and the actual pronunciation. The criterion frequency of occurrence clearly favors shorter N-grams, and it is therefore questionable whether this measure alone can provide satisfactory results. The criterion articulation rate does not seem to be very effective in identifying potential MWE candidates. Taken together these findings suggest that to identify MWEs an ingenious combination of criteria will probably be required.

Another problem that emerged from this study is that of the parent / child relationship between N-grams. By applying the selection measures we identified many N-grams that all seem to be related to some common core, e.g. many different N-grams were selected containing the word "gegeven", which are all related to the MWE "op een gegeven moment". In fact, these results are one of the reasons why we present the results for the case study on this MWE "op een gegeven moment" here. The question that arises at this point is how these N-grams should be treated. It is clear that we are dealing with sequences that can be problematic because they display a considerable amount of reduction. This would argue for treating such sequences as MWEs. On the other hand, it may seem bizarre to consider all these sequences with "gegeven" as MWEs because they do not seem to constitute a unified entity in any linguistic way, while this is the case for "op een gegeven moment". It stands to reason that while these considerations can easily be made by using our knowledge of the language, they are difficult to implement for automatic detection.

In general, the results presented in this paper reveal that the problem of pronunciation modeling of multiword expressions in spontaneous speech is indeed a real one. Once MWEs have been identified, it has to be decided how to handle them so as to model the huge variation in pronunciation that they display. Since the studies that have addressed pronunciation modeling of MWEs so far were limited to small numbers of MWEs, it is necessary to experiment with larger lexicons and higher numbers of MWEs to determine which approach is most promising. Another possibility would be to carry out evaluation within the application in which the multiword expressions are going to be used [23], be it automatic speech recognition, automatic phonetic transcription or speech synthesis.

To conclude, the study reported in this paper has shown that the problem of pronunciation variation in speech technology is probably more serious than one might have thought so far. Not only is there variation within words and across words, but words that often occur in combination may enter into such special relationships as to display pronunciation forms that never appear in other contexts. These new pronunciations can be so deviant as to require special treatment. Therefore, the most important conclusion of our study is that further research is needed e.g., to study alternative, automatic measures of deviant pronunciation, to determine the optimal procedure for MWE extraction, and to study how to handle MWEs in various applications (ASR, speech synthesis, and automatic phonetic transcription).

7. REFERENCES

The references are listed in alphabetic order.

[1] A. Bell, D. Jurafsky, E. Fosler-Lussier, C. Girand, M.L. Gregory, D. Gildea, D., "Effects of disfluencies, predictability, and utterance position on word form variation in English conversation," Journal of the Acoustical Society of America 113 (2), pp. 1001-1024, 2003.

[2] K. Beulen, S. Ortmanns, A. Eiden, S. Martin, L. Welling, J. Overmann, "Pronunciation modelling in the RWTH large vocabulary speech recognizer," Proceedings of the ESCA Workshop "Modeling Pronunciation Variation for Automatic Speech Recognition", Rolduc, Kerkrade, pp. 13-16, 1998.

[3] D. Binnenpoorte, C. Cucchiarini, L. Boves, H. Strik, "Multiword Expressions in Spoken Language: an exploratory study on pronunciation variation," Computer, Speech & Language 19(4), pp. 433-449, 2005.

[4] A. Elffers, C. Van Bael, H. Strik, "Adapt: Algorithm for Dynamic Alignment of Phonetic Transcriptions," CLST internal report, 2005.

[5] J.L. Fidelholtz, "Word frequency and vowel reduction in English," CLS 11, pp. 200-213, 1975.

[6] M. Finke, A. Waibel, "Speaking mode dependent pronunciation modeling in large vocabulary conversational speech recognition," Proceedings of EuroSpeech-97, Rhodes, pp. 2379-2382, 1997.

[7] E. Fosler-Lussier, N. Morgan, "Effects of Speaking Rate and Word Frequency on Pronunciations in Conversational Speech," Speech Communication 29: 137-158, 1999.

[8] M.L. Gregory, W.D. Raymond, A. Bell, E. Fosler-Lussier, D. Jurafsky, "The effects of collocational strength and contextual predictability in lexical production," Chicago Linguistics Society (CLS-99), pp. 151-166, 1999.

[9] D. Jurafsky, A. Bell, M.L. Gregory, W.D. Raymond, "Probabilistic Relations between Words: Evidence from Reduction in Lexical Production," In J. Bybee, P. Hopper, (eds.): Frequency and the emergence of linguistic structure, John Benjamins, Amsterdam, pp. 229-254, 2001.

[10] J.M. Kessens, M. Wester, H. Strik, "Improving the Performance of a Dutch CSR by Modeling Within-word and Cross-word Pronunciation Variation," Speech Communication 29: 193-207, 1999.

[11] C.H.A. Koster, "Transducing Text to Multiword Units," Proceedings MEMURA 2004 workshop, Lisbon, pp. 31-38, 2004.

[12] K. Kuiper "Smooth talkers. The linguistic performance of auctioneers and sportscasters," Mawah,NJ: Lawrence Erlabum Associates, 1996.

[13] J. Nivre, J. Nilsson, "Multiword Units in Syntactic Parsing," Proceedings MEMURA 2004 workshop, Lisbon, pp. 39-46, 2004.

[14] J. Odijk, "Reusable Lexical Representations for Idioms," Proceedings LREC 2004, Lisbon, pp. 903-906, 2004.

[15] N. Oostdijk, "The Spoken Dutch Corpus. Overview and first evaluation," Proceedings LREC-2000, pp. 887-894, 2000.

[16] N. Oostdijk, "The design of the Spoken Dutch Corpus," In: Peters P., Collins P., Smith A. (Eds.): New Frontiers of Corpus Research, Rodopi, Amsterdam, pp. 105-112, 2002.

[17] M. Pluymaeckers, "Prefabs in sports commentary," Master's thesis, Tilburg University, 2003.

[18] I.A. Sag, T. Baldwin, F. Bond, A. Copestake, D. Flickinger, "Multiword expressions: A pain in the neck for NLP," LinGO Working Paper (2001-03), http://lingo.stanford.edu/pubs/WP2001-03.ps.gz, 2001.

[19] T. Sloboda, A. Waibel, "Dictionary Learning for Spontaneous Speech Recognition," Proceedings of ICSLP-96:, Philadelphia, pp. 2328-2331, 1996.

[20] H. Strik, "Pronunciation adaptation at the lexical level," In: J-C. Juncqua, C. Wellekens (Eds.) Proc. of the ITRW 'Adaptation Methods For Speech Recognition', Sophia-Antipolis, France: 123-131, 2001.

[21] H. Strik, D. Binnenpoorte, C. Cucchiarini, "Multiword Expressions in Spontaneous Speech: Do we really speak like that?," Proc. of InterSpeech 2005, Lisbon, pp. 1161-1164, 2005.

[22] H. Strik, C. Cucchiarini, "Modeling pronunciation variation for ASR: a survey of the literature," Speech Communication, Vol. 29 (2-4), pp. 225-246, 1999.

[23] C. Van Bael, H. Strik, H. van den Heuvel, "Application-oriented validation of phonetic transcriptions: preliminary results," Proceedings of 15th ICPhS, Barcelona, Spain, pp. 1161-1164, 2003.

[24] G.K. Zipf, "Relative frequency as a determinant of phonetic change," Harvard Studies in Classical Philology 15: 1-95, 1929.

[25] http://lands.let.ru.nl/cgn/ehome.htm, CGN website

[26] http://www.phon.ucl.ac.uk/home/sampa/dutch.htm, SAMPA website