

# Durational Information in Word-initial Lexical Embeddings in Spoken Dutch

Odette Scharenborg

Centre for Language Studies, Radboud University Nijmegen, The Netherlands  
Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

o.scharenborg@let.ru.nl

## Abstract

There is a growing body of research showing the importance of durational information for the disambiguation of temporarily ambiguous speech due to lexical embedding (e.g., *rye* in *rises*) in laboratory settings. The current research investigates whether durational differences are present in non-laboratory speech. We focus on two types of speech: read speech and speech taken from interviews. Durations of thousands of instances of monosyllabic words and the same phonemic string embedded as the first syllable of a polysyllabic word (so-called embedded words) were obtained from the Spoken Dutch Corpus. These durations were first adjusted to many known sources of durational differences. A subsequent statistical analysis on these adjusted durations showed a significant difference in durations between monosyllabic words and embedded words for both speaking styles, suggesting that the presence of durational differences between monosyllabic words and embedded words is a general characteristic of spoken Dutch. Although the differences are small, it is argued that these durational differences are perceptually relevant.

**Index Terms:** durational information, ambiguity resolution, speech.

## 1. Introduction

While languages consist of tens of thousands of words, the phonemes of which they are made up are far fewer in number, typically around 25-30 phonemes [1]. Moreover, words can start and end at any time. Inevitably, many words resemble one another, and will (partially) overlap. For instance, the spoken utterance “the sun rises” /θəsʌnrʌɪzɪz/ not only contains the intended words *the*, *sun*, and *rises* but also embedded words, e.g., *I* and *rye*, and words that straddle word boundaries, e.g., *sunrises*. This temporary ambiguity, however, is usually quickly solved by listeners.

During spoken-word recognition, the number of activated words has been shown to influence word recognition times, with slower word recognition times with increasing number of activated words [2]. The extent of lexical embedding in a language can thus play a significant role during human spoken-word recognition [3]. Several studies have investigated the degree to which lexical embedding occurs, most of them on the basis of dictionaries (for English: [4]; Dutch: [5,6]). These studies showed that a majority of polysyllabic words have shorter words embedded in them, and that these words are most likely to be embedded word-initially. A few studies have extended this type of investigation into the realm of spoken language, investigating the question how often listeners encounter embedded words when listening. An analysis of lexical embedding on a real-speech corpus for English showed that over 70% of all polysyllabic word tokens contained at least one embedded word, and again the majority

of these embedded words appeared word-initially [7]. An analysis of lexical embedding in two corpora of spoken Dutch however showed a much lower occurrence of lexical embedding than the dictionary studies and the English real-speech corpus; on average only 12 to 20% (depending on the speech style) of the polysyllabic words had a word-initially embedded word [6]. Nevertheless, the frequency of word-initial lexical embedding in spoken language is a frequent phenomenon that listeners have to deal with on a daily basis.

Studies in phonetics and psycholinguistics give insights into how listeners are able to resolve this temporary ambiguity due to lexical embedding. Phonetic studies showed that for many languages the duration of a syllable decreases with increasing number of syllables in the word [8-11] (an exception is, e.g., Italian [12]). Moreover, there is now a vast amount of evidence from psycholinguistics [13-18] that listeners can use subtle phonetic information for the disambiguation of temporarily ambiguous stretches of speech. For instance, Salverda and colleagues concluded that the lexical interpretation of temporarily ambiguous sequences is influenced by duration, such that a longer sequence tends to be interpreted as a monosyllabic word more often than a shorter sequence [14]. In laboratory conditions, durational differences are thus found to exist between syllables in longer and shorter words; moreover, the human speech recognition system is able to use this durational information to disambiguate temporarily ambiguous speech due to lexical embedding.

The question addressed in this study is whether these durational differences between syllables in shorter and longer words exist in non-laboratory speech. In this study, we focus on word-initial lexical embedding, since this is by far the most frequent type of lexical embedding; moreover speech recognition is considered to operate from ‘left to right’ without backtracking. Specifically, this study investigates whether monosyllabic words and the identical phoneme string produced as the first syllable of a polysyllabic word (hereafter referred to as ‘embedded word’) differ significantly in duration in non-laboratory speech. The presence of durational differences between monosyllabic and embedded words is investigated for a speaking style that produces more carefully produced speech, i.e., read speech, and a speaking style that we may assume produces somewhat less carefully produced speech, i.e., interviews. These analyses will provide the first evidence of the occurrence and nature of durational information in the context of lexical embedding in non-laboratory spoken Dutch.

## 2. Materials

The data used in this experiment was taken from a previous study that investigated the frequency of occurrence of word-initial lexical embeddings in spoken Dutch [6], more specifically in the read speech and interview parts of the Spoken Dutch Corpus [19]. The Spoken Dutch Corpus is a corpus of almost nine million words of Dutch spoken in the

Netherlands and in Flanders (Belgium), in over 14 different speech styles or components, ranging from formal to informal.

Many factors are known to influence segment duration, including stress, position of the segment in a syllable and the position of the syllable in the word (see e.g., [20]). To control for as many of these factors, they were either kept constant or were entered as control variables in the statistical analyses (see Section 3). To that end, embedded words always appeared at the onset of the polysyllabic words. Moreover, monosyllabic and embedded words consisted of identical phoneme strings, lexical stress patterns, and syllable structure, e.g., *thee* (/te/; E: *tea*) in *tegen* (/te.xən/; E: *against*). Both the monosyllabic and embedded words appeared in the dataset, i.e., both *thee* and *tegen* had to exist in the dataset. In order to investigate the presence of durational differences between the monosyllabic and embedded words, manually verified phonemic transcriptions and segmentations of the speech files were available for all monosyllabic and embedded words. These segmentations were created using a forced alignment procedure identical to that in [21]. The set-up of the forced alignment procedure resulted in phoneme segmentations with a minimum phoneme length of 15 ms. These segmentations were used as the basis of this investigation.

Table 1 shows the number of types and tokens of the monosyllabic and embedded words for the read speech and interview components. The read speech set consists of 15 281 words produced by 133 speakers (with on average 114.9 words per speaker and 131 speakers with more than 10 words in the dataset). The interview set consisted of 3 547 words produced by 65 speakers with on average 54.6 words per speaker (and 43 speakers with more than 10 words in the set). Durations (in seconds) were computed for the monosyllabic words and the embedded words indicated in Table 1 on the basis of phoneme segmentations that were automatically derived and manually verified.

Table 1. Frequency counts of the monosyllabic and embedded words in the read speech and interview component of the Spoken Dutch Corpus used in the analyses.

	#monosyllabic words		#embedded words	
	Types	Tokens	Types	Tokens
Read speech	487	12 657	376	2 624
Interviews	179	3 211	93	336

### 3. Methodology

The presence of durational differences between monosyllabic and embedded words was investigated by statistically comparing the durations of the monosyllabic words with those of the identical phoneme string embedded as the first syllable of a polysyllabic word. This investigation was carried out for read speech and interview speech separately. All analyses were carried out using generalised linear mixed-effect models (e.g., [22]), containing both fixed and random effects. We used treatment coding for all categorical predictors.

For each of the speech styles a best-fitting model was built by including several fixed (see Section 3.1) and random predictors of word duration. The random effects included in the model were the orthographic forms of the monosyllabic/embedded words (these are necessarily identical) and the speakers. We started by building the most complex model, i.e., the model with all possible two-way and three-way interactions between the predictors. Subsequently, interactions and predictors that proved not significant were step-by-step removed from the model. The question of whether there are

durational differences between monosyllabic and embedded words is then answered by adding a factor called *Item\_type*, indicating whether the item is a monosyllabic or an embedded word, to the best-fitting model. The models with and without the additional factor are then compared by means of likelihood ratio tests (using the *anova* function in R).

Inspection of the distributions of the word durations showed that the durations of the monosyllabic and embedded words were not normally distributed. Therefore, the word durations were first transformed using a square root (*sqrt*) transform for read speech and a log transform for interviews. Figure 1 shows the distributions of the square root/log durations of the embedded words (solid lines) and the monosyllabic words (dashed lines) for the read speech and interview components. It is clear from Figure 1 that the transformed distributions of the durations of the embedded words and the monosyllabic words are largely overlapping, although different in shape.

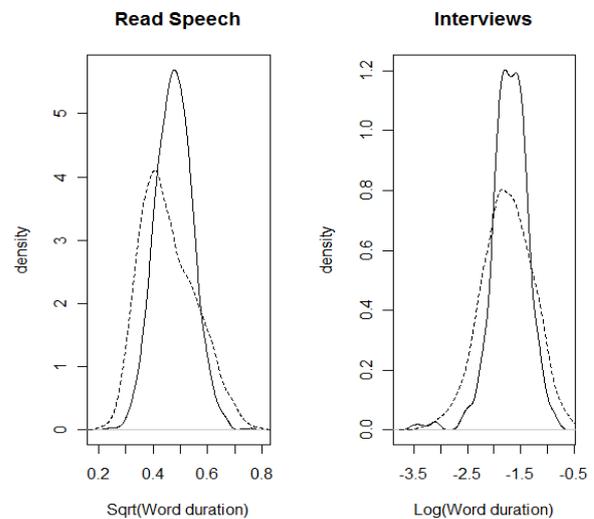


Figure 1. The distributions of the square root/log durations of the embedded words (solid lines) and the monosyllabic words (dashed lines) for the read speech component (left panel) and the interview component (right panel).

#### 3.1. Step 1

To further control for the influence of different duration factors, as many of these were added as control variables in the statistical analyses. These include segment identity and the position of the word in the phrase (note: due to the nature of our data we used position in the ‘chunk’ instead of phrase) [20]. Note that an important predictor of syllable duration, i.e., the number of syllables [20] was not included in our analyses because its heavily skewed distribution in our data set would effectively lead to a factor with two levels: the monosyllabic and the polysyllabic word, i.e., the distinction under investigation in this study. Other factors include the number of words in the chunk, speaking rate, and vowel duration. Moreover, a major contributor to durational differences is prosody [11]. In the absence of a prosodic description of the speech material, several part-of-speech (POS) variables were defined that capture important parts of prosodic durational variation (see Table 2). The factor POS of the word indicates whether the word is a content or function word. The factors POS tag of the preceding word and POS tag of the following

word give word status information of the immediate context of the word. Moreover, if the word is preceded or followed by a pause or if it is the start or end of a chunk, then the POS tag of the preceding or following word is scored as PAUSE. Thus, prosodic information concerning whether a word is preceded or followed by a pause is also included in these two variables.

Table 2 lists the variables and their type (including the levels in the case of categorical variables) that were used in the statistical analyses to control for these factors. The distributions of the continuous variables were investigated and transformed when not normal. Table 3 shows the descriptive statistics for the continuous variables and the applied transformations for each speaking style separately. Normalised position of the word in the chunk (i.e., a speech fragment of usually 2-3 s used for the automatic alignment of the spoken data) is calculated as the position of the word in the chunk divided by the number of words in the chunk. Speaking rate was calculated as the number of phonemes in the chunk divided by the duration of the chunk in ms. Inspection of the distribution of the categorical variables showed that all were highly skewed. All categorical variables were therefore reorganised into the levels listed in Table 2.

### 3.2. Step 2

Table 4 shows the correlations between the continuous variables for the read speech and interview components separately. The highest correlations for both speaking styles were found for speaking rate and vowel duration: For higher speaking rates, the (log/square root) vowel duration is shorter. All but one correlation (indicated with <sup>+</sup>) were significant at the  $p < .001$  level (after Bonferroni correction). For all significant correlations, residuals were created which were subsequently used in the statistical analyses.

### 3.3. Step 3

Each of the factors in Table 2 was entered as a fixed predictor to the model and investigated in all possible 2-way and 3-way interactions. When a predictor or a 2-way or 3-way interaction did not reliably improve model fit, it was removed. The final, best-fitting model, for each speech style separately, only contained predictors and interactions that were significant. In the two best-fitting models (hereafter referred to as Model 0), all  $t$ 's for the highest level interactions (or at least one of the contrasts in these interactions) were higher than 2.48.

Subsequently, for each speech style, Model 0 was used to create the residual of duration resulting in the duration that cannot be predicted on the basis of the fixed variables. This procedure ensures that if we find a difference in duration between monosyllabic and embedded words, it can indeed be attributed to the difference in monosyllabic vs. embedded word. In the final step, the predictor Item\_type (i.e., monosyllabic vs. embedded word) was added to the two Models 0 to answer the question whether there is a durational difference between monosyllabic words and embedded words.

## 4. Results

The mean duration of the monosyllabic words in the read speech component was 219.1 ms ( $SD = 99.4$  ms) and that of the embedded words was 232.4 ms ( $SD = 66.4$  ms). Contrary to what has been found in the literature, the duration of the embedded words was longer than those of the monosyllabic words, but note that these are the raw durations. These averages are based on durations that have not yet been normalised for the various factors that may influence duration.

Table 2. Overview of the variables and their type and levels (for categorical variables only) used in the analyses.

Description of variables	Type
number of words in the chunk	continuous
normalised word position in chunk	continuous
vowel duration	continuous
speaking rate	continuous
#phonemes in word	2 phonemes, 3+ phonemes
vowel identity	short, long
POS of the word	content, function
POS of the following word	content, function, pause
POS of the preceding word	content, function, pause

Table 3. Descriptive statistics for the continuous predictors for each speaking style separately (the applied transform).

Continuous predictors	Read speech		Interviews	
	mean	SD	mean	SD
Sentence_length (square root)	17.3	10.5	19.7	14.1
Normalised_position	0.5	0.3	0.5	0.3
Vowel_duration (Interviews: log transform; Read speech: square root)	0.11	0.05	0.10	0.06
Speaking_rate	11.6	2.0	12.3	3.2

Table 4. Correlations between the continuous factors. All correlations  $p < 0.001$ , apart from <sup>+</sup> ( $p > 0.1$ ).

Read speech	sqrt(Vowel_duration)	sqrt(Sentence_length)	Normalised_position
Speaking_rate	-0.35	0.17	-0.04
Normalised_position	0.08	-0.13	
Sentence_length	-0.05		
Interviews	log(Vowel_duration)	sqrt(Sentence_length)	Normalised_position
Speaking_rate	-0.32	0.20	-0.12
Normalised_position	0.08	-0.24	
Sentence_length	-0.03 <sup>+</sup>		

First, the best-fitting model (Model 0) was built to create the residual of duration (Intercept:  $\beta = .1725$ ,  $SE = .0061$ ,  $t = 28.08$ ,  $p < .001$ ). Subsequently, to investigate whether there are durational differences between monosyllabic words and embedded words in read speech, Item\_type was added to this model. The resulting Model 1 (Intercept:  $\beta = .1769$ ,  $SE = .0061$ ,  $t = 28.93$ ,  $p < .001$ ) showed a main effect for Item\_type ( $\beta = -.0135$ ,  $SE = .0013$ ,  $t = -10.05$ ,  $p < .001$ ;  $p$ -values based on Monte Carlo Markov Chain sampling), i.e., after residualising duration for all factors that might influence word duration, there is a significant difference in duration between embedded words and monosyllabic words such that embedded words have shorter durations than monosyllabic words. This is in line with results found in the literature [8-11]. Note, that the effect of all the other predictors and interactions did not change after adding Item\_type to the model.

In the final step, Model 0 and Model 1 were compared using the likelihood ratio test. Table 5 shows the results of this comparison (for the read speech and interviews component, separately). The comparison showed a significant improvement of Model 1 over Model 0. So, despite the overlap of the distributions for monosyllabic and embedded words shown in Figure 1, there is a significant durational difference between embedded and monosyllabic words in read speech.

Table 5. Evaluation results for the comparison of the best-fitting mixed-effect models without (Model 0) and with (Model 1) *Item\_type* for the read speech and interview components.

	Read speech		Interviews	
	Model 0	Model 1	Model 0	Model 1
Log likelihood	26921	26970	-239.5	-234.8
$\chi^2$ (compared to Model 0)		97.16		9.45
$p$ (compared to Model 0)		0.001		0.005

For the interview component, monosyllabic words were on average 195.4 ms long ( $SD = 98.3$  ms), whereas the average duration of embedded words was 192.3 ms ( $SD = 61.8$  ms). Durations were therefore on average shorter in the interview subset than in read speech, which is not surprising given the higher speaking rates for the interview subset (see Table 3). Again, first the best-fitting model (Model 0) was built to create the residual of duration (Intercept:  $\beta = -.4738$ ,  $SE = .0916$ ,  $t = -5.174$ ,  $p < .001$ ). Adding *Item\_type* to build Model 1 showed a significant effect of *Item\_type* ( $\beta = -.0697$ ,  $SE = .0229$ ,  $t = -3.045$ ,  $p < .005$ ; Intercept:  $\beta = -.4321$ ,  $SE = .0924$ ,  $t = -4.674$ ,  $p < .001$ ), again with shorter durations for embedded words than for monosyllabic words. The effect of all the other predictors and interactions did not change after adding *Item\_type* to the model. The comparison of Model 0 and Model 1 using the likelihood ratio test (see Table 5) showed that also in more spontaneous speech, as captured in the interviews, significant durational differences exist between embedded words and monosyllabic words.

## 5. General Discussion and Conclusion

This work was inspired by a body of research showing the role of durational information for the disambiguation of temporarily ambiguous speech due to lexical embedding. The research question addressed in this study is whether durational differences, which have repeatedly been shown to be important for the disambiguation of temporarily ambiguous speech in laboratory settings, exist between monosyllabic words and the identical phoneme string produced as the first syllable of a polysyllabic word in non-laboratory speech. Two speaking styles were investigated: read speech and speech obtained from interviews.

Durations of thousands of instances of monosyllabic words and the same phonemic string embedded as the first syllable of a polysyllabic word (so-called embedded words) were obtained from the Spoken Dutch Corpus. These durations were adjusted to many known sources of durational differences. One could argue that this set-up presents a worst case scenario, i.e., of a listener who is exposed to the speech of a person unknown to him. This is in contrast to a set-up where specific, identical syllables produced by the same speaker in monosyllabic and embedded contexts would be compared. Such an analysis would provide a better control for the extraneous factors we tried to accommodate within our statistical model. This is a topic for further study.

Results of the statistical analyses showed that significant differences in duration exist between monosyllabic and embedded words in both the read speech and the speech from interviews, suggesting that the presence of durational differences between monosyllabic words and embedded words is a general characteristic of spoken Dutch. Potentially, these significant durational differences between monosyllabic and embedded words can be used for resolving temporarily ambiguous phoneme strings in everyday speech. It is however

not evident that listeners are able to use these durational differences. As is clearly depicted in Figure 1: the duration distributions of the monosyllabic and embedded words are largely overlapping. Moreover, the effect sizes are rather small. A comparison of the intercepts of Model 0 and Model 1 for both speech styles shows that the difference between the monosyllabic and embedded words is about 2.5% for read speech and about 9% for the interviews.

For segments, which usually range in duration from 30 to 300 ms, differences in segment durations have to range from 10 to 40 ms for listeners to be able to hear these differences [8]. Roughly speaking, just-noticeable-differences (JNDs) range from 3% to 33% for segment durations. Quené investigated how large within-speaker variations in speaking rate have to be in order for these to be perceptually relevant. His study on speech from the same corpus used in this study, suggested that changes in speech tempo have to be larger than 10% to be noticeable [23]. Given these results, the durational differences found in the current study for read speech are most likely below the JND threshold for duration. Potentially, the durational differences found for interviews are above the JND threshold.

However, the body of research highlighting the role of durational information for the disambiguation of temporarily ambiguous speech clearly indicates that listeners do not need to be able to notice the durational differences in order to be able to use the durational differences for disambiguation. This is corroborated by results from two experiments investigating French liaison. [24] found that word recognition by French listeners was influenced by durational differences of only 12 ms of the [ʁ] in the two homophone word sequences *dernier oignon* and *dernier rognon*. However, when French listeners were asked to explicitly decide which word sequence they had heard, this did not result in identification accuracies above chance [25], even though the durational differences in that particular experiment were slightly higher than those of [24], i.e., 17 ms.

In Dutch, segmentations of temporarily ambiguous two item sequences such as *een(s)pilj...*, where both *een* (E: *a*) and *eens* (E: *once*) are Dutch words, and where both *pilj* and *spilj* are onsets of possible words (e.g., *pilj*, E: *pipe*; *spilj*, E: *nail*) are modulated by differences in /s/ duration of only 30 ms (about 3%) [17]. Dutch listeners thus might not be able to notice the small but significant durational differences found in this study, but since *noticing* is not crucial for *using* durational information for disambiguation of temporary ambiguous speech, it is likely that the found durational differences in the current study are still perceptually relevant. It thus seems likely that the laboratory findings on the role of durational information in disambiguation can be extended to non-laboratory speech and speech processing conditions.

## 6. Acknowledgements

This work is a follow-up of the MA project carried out by Stefanie Okolowski (University of Trier, Germany) under the supervision of the author at the Radboud University. The author would like to thank Barbara Schuppler and Iris Hanique for kindly providing the forced alignments, Roeland van Hout for discussions on the statistical analyses, and the members of the Sound Learning group of James McQueen for valuable comments on this work. This research was partly sponsored by a Vidi-grant from the Netherlands Organisation for Scientific Research (NWO; grant number: 276-89-003).

## 7. References

- [1] Maddieson, I., *Patterns of Sounds*, Cambridge University Press, Cambridge, 1984.
- [2] Norris, D., McQueen, J. M., & Cutler, A., "Competition and segmentation in spoken word recognition", *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21(1):1209-1228, 1995.
- [3] Cutler, A., Norris, D., & Sebastián-Gallés, N., "Phonemic repertoire and similarity within the vocabulary", *Proceedings of ICSLP, Korea*, 65-68, 2004.
- [4] McQueen, J., & Cutler, A., "Words within words: Lexical statistics and lexical access", *Proceedings of ICSLP, Canada*, 221-224, 1992.
- [5] Frauenfelder, U.H., "Lexical alignment and activation in spoken word recognition", in J. Sundberg, L. Nord, & R. Carlson [Eds.], *Music, Language, Speech & Brain*, 294-303, London: Macmillan, 1991.
- [6] Scharenborg, O., & Okolowski, S., "Lexical Embedding in Spoken Dutch", *Proceedings of Interspeech, Brighton, UK*, 1879-1882, 2009.
- [7] Cutler, A., McQueen, J., Baayen, H., & Drexler, H., "Words within words in a real-speech corpus", *Proceedings of the Australian International Conference on Speech Science & Technology (SST), Australia*, 362-367, 1994.
- [8] Lehiste, I., *Suprasegmentals*. MIT Press, Cambridge, MA, 1970.
- [9] Nootboom, S. G., *Production and perception of vowel duration in Dutch: A study of durational properties of vowels in Dutch*. PhD thesis, University of Utrecht, The Netherlands, 1972.
- [10] Quené, H., "Durational cues for word segmentation in Dutch", *Journal of Phonetics*, 20:331-350, 1992.
- [11] White, L., & Turk, A. E., "English words on the Procrustean bed: Polysyllabic shortening reconsidered", *Journal of Phonetics*, 38:459-471, 2010.
- [12] D'Imperio, M., Rosenthal, S., "Phonetics and Phonology of Main Stress in Italian", *Phonology*, 16(1):1-27, 1999.
- [13] Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G., "Leading up the lexical garden-path: Segmentation and ambiguity in spoken word recognition", *Journal of Experimental Psychology: Human Perception and Performance*, 28:218-244, 2002.
- [14] Salverda, A. P., Dahan, D., & McQueen, J. M., "The role of prosodic boundaries in the resolution of lexical embedding in speech comprehension", *Cognition*, 90:51-89, 2003.
- [15] Cho, T., McQueen, J. M., & Cox, E. A., "Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English", *Journal of Phonetics*, 35:210-243, 2007.
- [16] Andruski, J. E., Blumstein, S. E., & Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition*, 52, 163-187.
- [17] Shatzman, K. B., & McQueen, J. M., "The modulation of lexical competition by segment duration", *Psychonomic Bulletin & Review*, 13:966-971, 2006.
- [18] Blazej, L. J., Cohen-Goldberg, A. M., "Can we hear morphological complexity before words are complex?" *J of Experimental Psychology: Human Perception and Performance*, 41(1):50-68, <http://dx.doi.org/10.1037/a0038509>, 2015.
- [19] Oostdijk, N. H. J., Goedertier, W., Van Eynde, F., Boves, L., Martens, J.-P., Moortgat, M., & Baayen, H., "Experiences from the Spoken Dutch Corpus project", *Proceedings of LREC, Las Palmas de Gran Canaria*, 340-347, 2002.
- [20] van Santen, J., "Contextual effects on vowel duration", *Speech Communication*, 11:513-546, 1992.
- [21] Schuppler, B., Ernestus, M., Scharenborg, O., Boves, L., "Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions", *Journal of Phonetics*, 39:96-109, 2011.
- [22] Baayen, R. H., Davidson, D. J., & Bates, D. M., "Mixed-effects modeling with crossed random effects for subjects and items", *Journal of Memory and Language*, 59:390-412, 2008.
- [23] Quené, H., "What is the Just Noticeable Difference for tempo in speech?", in H. Quené & V. van Heuven [Eds], *On Speech and Language, Studies for Sieb G. Nootboom*, 149-158, 2004.
- [24] Spinelli, E., McQueen, J. M., Cutler, A. (2003), "Processing resyllabified words in French", *J of Memory and Language*, 48(2), 233-254. doi:10.1016/S0749-596X(02)00513-2, 2003.
- [25] Shoemaker, E., Birdsong, D., "La resolution de la liaison en français par des locuteurs natifs et non-natifs", *Acquisition et Interaction en Langue Étrangère*, 27:43-62, 2008.