

Advanced second language learners experience difficulties processing reduced word pronunciation variants

Abstract

Words are often pronounced with fewer segments in casual conversations than in formal speech. Previous research has shown that foreign language learners and beginning second language learners experience problems processing reduced speech. We examined whether this also holds for advanced second language learners. We designed a dictation task in Dutch consisting of sentences spliced from casual conversations and an unreduced counterpart of this task, with the same sentences carefully articulated by the same speaker. Advanced second language learners of Dutch produced substantially more transcription errors for the reduced than for the unreduced sentences. These errors made the sentences incomprehensible or led to non-intended meanings. The learners often did not rely on the semantic and syntactic information in the sentence or on the subsegmental cues to overcome the reductions. Hence, advanced second language learners also appear to suffer from the reduced pronunciation variants of words that are abundant in everyday conversations.

Keywords: reduced word pronunciation variants; second language acquisition; word recognition; Dutch; casual speech; dictation task

1. Introduction

In many languages, words are often pronounced differently in casual conversations than in formal speech situations. In everyday conversations, words are frequently pronounced with weakened segments (e.g., a schwa instead of a full vowel) and with fewer segments or even with fewer syllables (see Ernestus & Warner, 2011 for an overview). For instance, in Dutch casual conversations, *Canada* [kanada] ‘Canada’ is often pronounced as [kanəda], *mist* [mist] ‘mist’ as [mɪs] and *wedstrijd* [vɛtstreit] ‘contest’ as [vɛs]. These reduced word pronunciation variants are easily understood by native listeners. This paper investigates how well reduced words are recognized by advanced second language learners.

Dutch is among the few languages for which the reduced pronunciation variants of words have been thoroughly investigated. Schuppler, Ernestus, Scharenborg and Boves (2011), for instance, reported that in the Ernestus Corpus of Spontaneous Dutch (Ernestus, 2000), 40% of the word tokens have weakly articulated or completely missing segments, while 19% percent of the word tokens lack syllabic nuclei. These figures are similar to those reported for American English (e.g., Johnson, 2004) and European French (e.g., Adda-Decker, Boula de Mareüil, Adda, & Lamel, 2005).

Some reduced word pronunciation variants deviate from the unreduced variants by just a few segments. These variants typically show regular reduction patterns, for instance, the reduction of a full vowel to schwa, the absence of a schwa, the absence of [l] or [r] after a vowel, or the absence of [t] after [s] (see e.g., Schuppler et al., 2011, for the description of some of these patterns in Dutch). Some tokens of words may simultaneously show several of these regular reduction patterns, and consequently deviate substantially from the words’ unreduced pronunciations, while still being highly

predictable. An example is the Dutch word combination *openbaar vervoer* /opəbarvərvur/ ‘public transportation’, which may be pronounced as /obavur/ due to schwa deletion and /r/ deletion in combination with voice assimilation and degemination.

In addition, some words have idiosyncratic extremely reduced pronunciation variants. Most of these are highly frequent words, for instance Dutch *daarom* [darɔm] ‘therefore’, which is often pronounced as [dam]. However, less frequent content words can also be extremely reduced. For instance, Dutch *wedstrijd* [vɛtstreit] ‘contest’ can be pronounced as [vɛs] (see Schuppler et al., 2011 for more Dutch examples).

Native listeners of a language typically have little difficulty understanding reduced word pronunciation variants. They are guided in their comprehension by the meaning (e.g., van de Ven, Tucker, & Ernestus, 2011) and syntactic structure (e.g., Tuinman, Mitterer, & Cutler, 2014; Viebahn, Ernestus, & McQueen, 2015) of the sentence, by the probability of the word given the surrounding words (e.g., van de Ven, Ernestus, & Schreuder, 2012), by speech rate (e.g., Dilley & Pitt, 2010), and by subtle characteristics of the speech signal (e.g., a relatively long [s] cueing a following absent [t], Zimmerer & Reetz, 2014). Furthermore, native listeners appear to have lexically stored at least some reduced word pronunciation variants together with their frequencies of occurrence (e.g., Ranbom & Connine, 2007), as they recognize a variant of a word more often if this variant occurs more often (see Ernestus, 2014, for a discussion).

In contrast, adult learners of a foreign language have difficulty recognizing words that are reduced. Nouveau (2012), for instance, found that in a dictation task with isolated words, Dutch advanced learners of French living in the Netherlands correctly transcribed no more than 56% of the French words that were produced without schwa

(e.g., *pelouse* [pəluz] ‘lawn’ pronounced as [pluz]). Wong and colleagues (2015) found that in an English dictation task, Hong Kong students with intermediate to high proficiency levels in English made errors for 40% of the words showing regular simple segment reduction. Finally, Brown & Hilferty (1986) reported that Chinese learners of English with an intermediate proficiency level correctly transcribed up to 35% of the English words and word combinations showing frequent weak reduction or contraction (e.g., *gonna* for *going to*) if they had not received any explicit training on these reduced variants. After attending 30 ten-minute lessons, they performed significantly better (68% correct), but still rather poorly.

It is likely that foreign language learners have problems understanding reduced word pronunciation variants because they have little experience with casual speech. Teachers tend to speak slowly and to hyper-articulate during language classes and the listening materials presented in class typically contain formal speech (e.g., read aloud speech or radio interviews). Foreign language learners thus do not hear many reduced word pronunciation variants.

In contrast, second language learners who live in the country of which they are learning the language are more likely to encounter these variants, and hence may experience less difficulty understanding them. They may have learnt to use the semantic, syntactic, acoustic, and frequency information that native listeners rely on.

Ten Bosch et al. (2016) reported on a dictation task with learners of Dutch as a second language. These learners had a low level of proficiency: they followed a course taking them from the A2 level to the B1 level of the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2011). The learners transcribed either sentences without reduced words or they transcribed the same sentences with

several words showing different types of reduction, ranging from one regular segment reduction to extreme, word idiosyncratic, reduction. The group of learners transcribing the reduced sentences were three times more likely to miss a syllable or word than the group of learners transcribing the unreduced sentences.

The low proficiency learners tested in Ten Bosch et al. (2016) may not have outperformed the foreign language learners because their level was too low to engage in casual conversations. Consequently, they probably had little experience with casual speech, like foreign language learners have. Therefore, in this study, we investigated whether advanced second language learners who follow a course that should bring them to the B2 CEFR level also have problems recognizing reduced words.

The dictation task used by Ten Bosch and colleagues (2016) was specifically developed for beginning learners of Dutch. The speech was read aloud, at a low speech rate, and the speaker consciously inserted reduced word pronunciation variants. This may have resulted in unnatural speech patterns. Because we tested learners at a higher proficiency level, our participants had a broader range of vocabulary and could process higher speech rates. We could therefore find spontaneously uttered sentences with reduced word pronunciation variants that we could incorporate in a dictation task. We extracted such sentences from a corpus of casual conversations, and used these to create the reduced version of our dictation task. The same speaker rerecorded the selected sentences without reductions, thus creating the unreduced version of the task. We examined the number and types of errors a group of learners made for the sentences with reduced words and the number of times they listened to these sentences, relative to how another group of learners performed on the unreduced versions of these sentences.

In addition, we tested how native listeners of Dutch perform on the two types of sentences.

2. Method

2.1. Participants

2.1.1. Second language learners

We tested 34 highly educated second language learners of Dutch, who were paid to participate in the experiment. They were recruited via Radboud in'to Languages (Nijmegen, the Netherlands), where they took a course that would bring them to the CEFR B2 level of Dutch. One of these participants did not understand the task and provided circumscriptions of the words to be transcribed (e.g., *de helft* 'half' instead of *50 procent* 'fifty percent') and we therefore excluded this participant from the dataset. The remaining 17 male and 16 female participants ranged in age from 19 - 53 years (mean: 32.6 years).

The participants were semi randomly assigned to one of the two versions of the dictation task such that the two groups represented similar native languages. The 16 participants transcribing the unreduced version of the experiment were native speakers of Arabic (three participants), English, Farsi (two participants), Kurdish, Polish, Portuguese, Romanian, Russian (two participants), Tagalog, Tigrinya, and Urdu. The 17 participants transcribing the reduced version were native speakers of Arabic (four participants), Armenian, Czech (two participants), German, Farsi (two participants), Italian (three participants), Kurdish (two participants), Ukrainian, and Mandarin.

2.1.2 Native listeners

The dictation task was also performed by 23 native listeners of Dutch: seven men and sixteen women, ranging in age from 18 - 27 years (mean age: 21.9 years). Twelve participants transcribed the reduced version of the task, while 11 participants transcribed the unreduced version. They were paid for their participation.

2.2. Materials

We selected the materials for the reduced version of the dictation task from the Ernestus Corpus of Spontaneous Dutch (Ernestus, 2000). This corpus contains casual conversations between pairs of highly educated, male native speakers of Dutch. Part of the corpus consists of free conversations while the other part consists of a role play with business negotiations. The recordings were made with a sampling frequency of 44.1 kHz.

We selected sentences from Speaker G, who was in his forties at the time of the recordings and had lived all his life in the western part of the Netherlands. He showed common reduction patterns while still speaking relatively clearly. We selected 29 tokens of different words and word combinations (hereafter called ‘items’) that occurred in contexts of which most words are known to learners of Dutch at the CEFR B1 level (see Table 1 for the list of target items). We preferred items that occurred in the same sentence or in neighbouring sentences or that occurred in contexts on the same topics. We spliced out the sentences with these target items, together with preceding or following sentences if these surrounding sentences were necessary for the sentences with the target items to make sense. We ordered these sentences in 14 groups with semantically and syntactically coherent content. Each sentence group contained between

one and four target items and was between 13 and 45 words long (mean 25.6 words). The Appendix provides the orthographic transcriptions of these sentence groups. They were paired in such a way that the second member of a pair formed a logical continuation to the first member.

Table 1. The target items in the dictation task

Sentence group	Orthographic transcription	Unreduced pronunciation	Pronunciation in dictation task	English translation
1	<i>dezelfde</i>	[døzɛlfdə]	[døzɛl]	‘the same’
	<i>als</i>	[als]	[as]	‘as’
2	<i>anders</i>	[andərs]	[andəs]	‘different’
	<i>precies</i>	[prəsis]	[pəsis]	‘exactly’
3	<i>zaterdagochtend</i>	[zatərdaxɔxtənt]	[satəɔxənt]	‘Saturday morning’
	<i>ik weet niet of</i>	[ɪk vɛt nit ɔf]	[gʋeniəf]	‘I don’t know if’
4	<i>alles is</i>	[aləs is]	[aləz'z]	‘everything is’
	<i>natuurlijk</i>	[natyrlək]	[natyk]	‘of course’
5	<i>gewoon</i>	[xəʋon]	[xʋon]	‘as usual’
6	<i>eigenlijk</i>	[ɛixələk]	[ɛig]	‘actually’
	<i>als het goed is</i>	[als ət xut is]	[sxuts]	‘if it is correct’
7	<i>verkopen</i>	[vɛrkopə]	[fkop]	‘sell’
	<i>bibliotheken</i>	[biblijotekə]	[bibltekə]	‘libraries’

8	<i>maken ze een</i>	[makə zə ən]	[maksən]	‘they make a’
9	<i>honderd</i>	[hɔndərt]	[hɔndəd]	‘hundred’
	<i>eind</i>	[ɛint]	[ɛin]	‘end’
	<i>dan kan ik er</i>	[dan kan ik ɛr]	[dɑŋ kɑŋkə]	‘then I can there’
10	<i>moet ik eens</i>	[mut ik ens]	[mudɪgəz]	‘should I once’
	<i>ik denk dat dat een</i>	[ik dɛŋk dat dat ən]	[dɛŋktatətən]	‘I think that that a’
11	<i>voor een deel</i>	[vor ən dɛl]	[fɔn dɛl]	‘for a part’
	<i>verkeerd</i>	[vɛrkert]	[fkɪt]	‘wrong’
12	<i>lekker warm</i>	[lɛkər vɑrm]	[lɛkəvɑrm]	‘nicely warm’
13	<i>jaren geleden</i>	[jarə xələdə]	[jarxlədə]	‘years ago’
	<i>waren we een keer</i>	[vɑrə və ən ker]	[vɑ və ker]	‘were we once’
	<i>electriciteit</i>	[elɛktrisiteit]	[elɪsiteit]	‘electricity’
	<i>vijftig procent</i>	[fɛiftəx prɔsɛnt]	[fɛiftprɔsɛnt]	‘fifty percent’
14	<i>realiseer</i>	[rɛjaliser]	[rɛjəsɛr]	‘realize’
	<i>Nederland</i>	[nedərlɑnt]	[nɛlɑnt]	‘The Netherlands’
	<i>absoluut</i>	[ɑpsɔlyt]	[ɑpsəl]	‘absolutely’

We asked the same speaker to record the sentences again, but without reduction. To ensure that the two versions of a sentence mainly differed in the presence versus absence of reduction, the speaker first listened to a sentence several times and then tried

to repeat it with the same speech rate and intonation, as well as with the same hesitations and interjections. These sentences were also recorded with a sampling frequency of 44.1 kHz. We manipulated the intensity of these recordings and the amount of background noise in order to match them to the original recordings. The reduced and unreduced versions of the sentence groups have approximately the same durations (mean duration of the reduced versions: 8.4 seconds; mean duration of the unreduced versions: 8.8 seconds; paired $t(13) = -1.49, p > 0.1$).

2.3. Experimental procedure

The participants performed the dictation task via the web and listened to the sentence groups on head phones. The learners were tested in three groups in a room with computers. The native listeners were tested individually in a sound-attenuated booth with one computer.

After the participants logged in, they provided some personal information (e.g., native language, age, number of years spent in the Netherlands). They then saw a screen informing them about the experiment. They read that they would hear fragments from spontaneous conversations and that they would see orthographic transcriptions of these fragments. The orthographic transcriptions would lack some words, indicated by a numbered line at the position of these missing words. The participants were asked to type in the missing words in the square below the orthographic transcription. We did not explicitly mention the hesitations in the fragments, which were transcribed as “uh” (see Figure 1 for an example), or the beep that replaced a name in a sentence, transcribed as *<naam>* ‘<name>’, because we believed that participants would easily understand the

meanings of these transcriptions themselves. The participants also read that they would hear the same speaker in all trials, discussing several topics.

Each sentence group was presented on a different page. We replaced not only the target items but also some more words and word combinations with horizontal lines. These additional words were easily intelligible and functioned as filler items.

INSERT FIGURE 1 AROUND HERE

Participants could listen to a sentence group as often as necessary by clicking on the repeat button. We assumed that this would increase participants' performance such that we would rather obtain an overestimation than an underestimation of how well they recognize reduced words. Moreover, this procedure decreases the effect of the participants' working memory spans on their performance, which may be an effect specific to dictation tasks. The program recorded the participants' transcriptions as well as how often they clicked on the repeat button. This provided us with two dependent variables that could show the effect of reduction on speech processing.

Before seeing the orthographic transcription of the first sentence group of a pair, the participants were given some summary information (e.g., *Het volgende fragment gaat over een radioprogramma over boeken*. 'The next excerpt is about a radio broadcast about books.'). Three pairs of sentence groups contained eight words in total that the learners may not know. These words were explained in the same way as new words are explained in the learners' textbooks, directly after the summary information about the sentence group pair (see Figure 2 for an example).

INSERT FIGURE 2 AROUND HERE

The experiment started with a practice trial. The participant listened to a sentence especially recorded for this purpose. They saw the orthographic transcription and the numbered horizontal lines replacing the two word combinations that had to be transcribed. By pressing a button, they could see these missing word combinations filled in the square below the orthographic transcription, each preceded by its number between square brackets.

The 14 sentence groups were presented in the same semantically most coherent order to every participant. The whole experimental session took a maximum of 20 minutes. After the experiment, the participants indicated whether their transcriptions could be used for scientific research.

2.4 Data analysis procedure

We scored participants' transcription accuracies for the 29 target items. We scored a transcription as correct in case it showed that the participant had understood what was said. We thus ignored spelling errors and the absence of determiners. On the other hand, wrong verb forms (e.g., first person singular present tense *verkoop* instead of plural present tense *verkopen* 'sell') or wrong derivational forms (e.g., *verkopers* 'sellers' instead of *verkopen* 'to sell') were scored as errors. Errors of these latter types, which need not result from misinterpretation due to reduction but may also be due to lack of grammatical knowledge, form a very small minority of the errors that we attested.

We analysed the number of times a participant listened to a sentence group and the number of errors a participant made for a sentence group with mixed effect

modelling. Our two predictors of interest were the version of the dictation task (reduced or unreduced), henceforth simply referred to as ‘reduction’, and the participant group (learners or native listeners). We included as random effects the participant and the sentence group.

For the analysis of the errors, we also tested statistical models with as additional predictor the number of times the participant has listened to the sentence group. Because the program did not register for every trial how often the participant listened to the sentence group (see below), these analyses were based on fewer trials than the models reported below. The number of repetitions showed a statistically significant effect in none of the statistical models.

2.5. Results

2.5.1. The data points for analysis

Due to technical problems, the experimental program did not record the transcriptions provided by four native listeners for Sentence group 1, and failed to record the numbers of times all participants listened to this sentence group. Further, for seven native listeners, it did not record the number of times these participants clicked on the repeat button at all.

The native listeners produced a total of 26 errors, of which six were in trials with unreduced speech. Errors mostly resulted from the participant not providing any transcription at all or transcribing a word (combination) that was already present in the transcription provided to them instead of the item that was absent. The natives produced relatively more errors for the reduced version of Sentence group 6 than for the unreduced and reduced versions of other sentence groups. This stretch of speech

contained a reduced word (*eigenlijk* ‘actually’) directly followed by a reduced word combination (*als het goed is* ‘if it is right’) with similar meanings. Four participants of 12 only transcribed the reduced word, while three others only transcribed the following reduced word combination. Because the reduced version of this sentence group is more difficult to transcribe than the reduced versions of the other sentence groups, we decided to leave both versions of this sentence group out of the further analyses of the data. Importantly, this decision does not affect the conclusions that we drew from the data, as was shown by additional analyses including Sentence group 6.

2.5.2. Analysis of the number of repetitions

We first compared the learners’ and the native listeners’ performances as to how often they listened to the reduced and the unreduced versions of the sentence groups. Because the program had not recorded this information for every trial, this analysis is only based on 637 trials (429 from the learners). On average, the learners listened 3.5 times to a sentence group, with a range of once to sixteen times. The native listeners seldom listened to sentence groups more often than once (only in 66 out of 208 trials).

The numerical data show that learners hearing the reduced target items listened more often to the sentence groups (on average 4.72 times) than the learners who heard the unreduced target items (on average 4.22 times). The same pattern holds for the native listeners (on average 1.74 versus 1.44 times). Because the Dutch listeners frequently only listened to a speech stretch once, the number of repetitions does not form a normal distribution. We therefore analysed the data by means of a binomial dependent variable indicating whether a given participant had listened to a given sentence group more often than once. We used generalized linear mixed effects models

with the logit link function. This model showed simple main effects for reduction (reduced: $\beta = 1.20$, $z = 2.324$, $p < 0.05$) and for participant group (native listeners: $\beta = -5.19$, $z = -6.948$, $p < 0.001$), as well as an interaction between these two variables (a native listener listening to unreduced speech: $\beta = -1.77$, $z = -2.159$, $p < 0.05$). No random slopes were significant. Learners listened more than once in nearly all trials: those who heard the reduced target items in 95% of the trials and those who heard the unreduced target items in 96% of the trials. In contrast, the Dutch natives who heard the reduced target items clicked on the repeat button significantly more often (in 43% of trials) than the Dutch natives who heard the unreduced target items (in 25% of trials).

We also performed an analysis just focusing on the learner data. We conducted linear mixed effect modelling with the number of repetitions as dependent variable. Reduction only emerged as a statistically significant random slope for sentence group, showing that reduction increased the number of times the learners listened to some sentence groups, while there was no effect (or a small negative effect) of reduction for other sentence groups (the estimate of the non-significant simple effect of reduction is 0.5018, while the random slope ranged between -1.0544 and 1.9055).

2.5.3. Quantitative analysis of the errors

We then compared the learners with the natives in the number of errors they produced. The learners produced 461 errors in 429 trials versus 18 errors produced by the natives in 295 trials. The numerical data show that the learners who heard reduced target items generally made more errors (347) than the learners who heard the unreduced versions (114). The natives produced a total of 13 errors for the reduced target items and five for the unreduced ones, a statistically non-significant difference.

Because the Dutch natives produced only few errors, the number of errors does not form a normal distribution. We therefore further analysed the errors produced by the natives and the learners by means of a binomial dependent variable, indicating whether a given participant produced minimally one error for a given sentence group. The Dutch natives hearing unreduced target items correctly transcribed 97% of the sentence groups, and those hearing the reduced versions 84% of the sentence groups. The learners hearing the unreduced target items correctly transcribed 62% of the sentence groups while those hearing the reduced versions only transcribed 30% of the sentence groups without errors. We used generalized linear mixed effects models, with the logit link function. We found a statistically significant simple main effect for participant group ($\beta = -4.5329$, $z = -9.267$, $p < 0.0001$): learners were more likely to produce errors than native listeners. More importantly, the interaction between reduction and participant group showed a marginally significant effect ($\beta = 1.2133$, $z = 1.709$, $p < 0.5$ one tailed; $p = 0.088$ two tailed): the effect of reduction is larger for learners than for native listeners.

We also performed an analysis on just the learners' data, investigating the effect of reduction on the total number of errors. That is, instead of analysing whether a sentence group was either transcribed correctly or with at least one error, we took every error into account. A linear mixed effect model with as dependent variable the number of errors in the sentence group showed a clear effect of reduction ($\beta = 1.0120$, $t = -4.925$, $p < 0.0001$).

2.5.4. Qualitative analysis of the errors

Finally, we performed a qualitative analysis of the transcription errors produced by the learners for the reduced target items. We focused on those items that were correctly transcribed by the majority of learners hearing the unreduced version of the dictation task. Like the natives, the learners sometimes did not provide transcriptions at all or they transcribed the wrong parts of the sentence group. The learners typically omitted reduced items that other learners transcribed incorrectly. For instance, while the word *verkeerd* [vɛrkert] ‘wrong’ reduced to [fkit] was ignored by six learners, it was transcribed incorrectly by four participants. Similarly, five learners hearing *Ik weet niet of* [ɪk vɛt nit ɔf] ‘I don’t know if’ reduced to [gʋeniəf] did not transcribe the phrase at all, while five transcribed the phrase incorrectly.

The learners often replaced target items by words with different meanings. Many of these replacements do not make sense in the context. Examples include the frequent replacement of reduced *natuurlijk* ‘of course’ by *toe* ‘to’, the replacement of reduced *bibliotheek* ‘libraries’ by *betekent* ‘means’, and the replacement of *electriciteit* ‘electricity’ by *feit* ‘fact’, *faciliteit* ‘facility’, *universiteit* ‘university’, and *tijd dat* ‘time which’. These results show that the learners did not fully take the semantic and syntactic cues in the context into account.

Other frequent replacements for reduced words fit the context, including *één december* ‘first of December’ for *eind December* ‘end of December’, and *vijf procent* ‘five percent’ for *vijftig procent* ‘fifty percent’. The acoustic differences between the unreduced variants of the intended words and of the incorrect transcriptions are small ([ɛint desɛmbər] versus [en desɛmbər] and [fɛiftəx prosɛnt] versus [fɛif prosɛnt]). For instance, unreduced *eind* and *één* differ in the quality of the vowel ([ɛi] versus [e]) and

in the presence versus absence of [t]. In the reduced sentence, *eind* was produced without clear [t]. The learners' transcription errors indicate that they could not benefit from the difference in vowel quality between *eind* and *één* or from the subsegmental cues to the /t/.

In contrast, some reduced word pronunciation variants showing simple segment reduction did not impose any problems. Among these are [ʌndəs] for [ʌndərs] *anders* 'other', [xʊʊn] for [xəʊʊn] *gewoon* 'as usual', [hʊndəd] for [hʊndrəd] *honderd* 'hundred' and [lɛkə] for [lɛkər] *lekker* 'nicely'. Perhaps more surprisingly, a few reduced word pronunciation variants that deviated substantially from the unreduced variants did not pose serious problems. These include [satəʊxənt] for [zatərdaxɔxtənt] *zaterdagochtend* 'Saturday morning', [nelant] for [nedərlant] *Nederland* 'the Netherlands', and [dəzɛl] for [dəzɛlfdə] *dezelfde* 'the same'. The position of *zaterdagochtend* in the sentence showed that the word was likely to be a time specification, and the position of *Nederland* as well as the preceding word *in* 'in' that this word was likely to be a place specification. Further, the words *dezelfde...als* form a highly frequent skip gram. These results thus suggest that the learners could rely on strong grammatical cues.

3. General Discussion

Previous studies have shown that foreign language learners, who mostly learn the language at school, have problems recognizing words when these occur in reduced pronunciation variants (e.g., Nouveau, 2012; Wong et al., 2015). This may not be the case for second language learners, who learn the language of their new country, because they are likely to have more experience with casual speech than foreign language

learners. Ten Bosch and colleagues (2016) showed, however, that second language learners with a low proficiency level also cannot recognize reduced words very well. We investigated how well reduced words are recognized by second language learners at an advanced level.

We conducted a dictation task with learners of Dutch who all followed the same course that should bring them to the CEFR level of B2. They were semi randomly assigned to two groups such that the two groups of learners were very similar in the native languages that they represent, containing similar numbers of native speakers of Arabic, Farsi, Kurdish, Roman languages, Slavic languages, and of languages closely related to Dutch. Because the two groups contained learners with similar backgrounds, they are not likely to differ much in their average Dutch proficiency. One listener group heard stretches of speech with reduced words, while the other group only heard unreduced words. The learners saw the orthographic transcriptions of the speech and had to fill in the missing words. They could listen to a speech stretch more than once.

Nearly all experiments investigating how non-native listeners process reduced word pronunciation variants tested read or acted speech. In contrast, the stretches of reduced speech that we tested were spliced from spontaneous conversations. The speech stretches with only unreduced words resulted from the same speaker shadowing his own speech stretches as closely as possible except for reduction. This choice of materials has the advantage that the reductions in our stimuli are completely natural and that the experiment is therefore ecologically more valid.

We first tested the dictation task with native listeners of Dutch. They made errors in 8% of the trials, mostly by not providing a transcription at all, or by providing the transcription of a word that was already present in the transcription provided. These

errors may have been the result of how we asked participants to enter their answers: in a square below the transcription of the speech stretch rather than directly in the transcription itself.

The Dutch natives hearing the reduced target items did not make significantly more errors than the Dutch natives hearing the unreduced versions. Nevertheless, reduction seems to have hindered the Dutch natives hearing the reduced target items as they listened significantly more often to their speech stretches than the Dutch natives hearing the unreduced items. This result is in line with previous studies showing that native listeners also experience more problems processing reduced word pronunciation variants than unreduced ones (e.g., Drijvers, Mulder & Ernestus, 2016 and references therein).

Concerning the learners, there was no statistically significant difference between the learners who heard the unreduced target items and those who heard the reduced target items in how often they listened to the speech stretches. Possibly, the learners hearing the reduced target items did not want to listen to the speech stretches more often because they already spent so much time on each trial or because they believed that listening to the speech stretches more often would not increase their accuracy. If they thought the latter, they were probably correct since the number of times learners listened to a speech stretch did not correlate with their performance on this speech stretch.

The learners hearing reduced target items produced many more errors (they made no errors for 30% of the sentence groups) than the learners transcribing the unreduced items (they made no errors for 62% of the sentence groups). Learners provided no transcriptions at all for many reduced target items. Since learners typically failed to provide transcriptions for items for which other learners provided incorrect

transcriptions, we believe that most learners' omissions are due to word recognition problems. In addition, the learners often replaced reduced items with similarly sounding words, probably also due to word recognition problems. These results strongly suggest that not only foreign language learners and second language learners of a low proficiency experience difficulty recognizing reduced words (e.g., Nouveau 2012, Ten Bosch et al., 2016): so do second language learners.

The learners especially produced many errors for the extremely reduced items, even for the high frequency ones like [natyk] for [natyrlək] *natuurlijk* 'of course'. In contrast, the learners made relatively few errors for reduced items showing regular simple segment reduction. Possibly they are able to overcome these simple reduction patterns because the same patterns also occur in their native languages (e.g., Mitterer & Tuinman 2012), or because they have learnt to accept small deviations from a word's unreduced pronunciation if this deviation does not result in another Dutch word they know.

The types of errors the learners made shed some light on why they experience problems processing reduced word pronunciation variants. First, we found that the learners did not always rely on the semantic and syntactic information in the context. This finding is in line with the results by Van de Ven, Tucker, and Ernestus (2010), for instance, who could not find semantic priming for second language learners.

Our results also suggest, however, that the learners relied on semantic and syntactic cues that were very strong. Possibly, the learners would have been better at recognizing the reduced words if they had heard more context, providing them with stronger semantic cues. If so, learners need more semantic context to recognize reduced words than native listeners do. This possibility could be addressed in future research.

The second type of information that the learners appeared not to make full use of were the details of the acoustic signal. For instance, the learners often misinterpreted reduced *eind December* ‘end of December’ as *één December* ‘first of December’. These transcription errors indicate that the learners could not benefit from the difference in vowel quality between *eind* and *één* or from the subsegmental cues to the /t/.

Note that the interpretation of *eind December* as *één December* may go unnoticed because both interpretations fit the context. Similarly, the frequent misinterpretation of *vijftig procent* ‘fifty percent’ as *vijf procent* ‘five percent’ may have gone unnoticed. Reduction may thus be more detrimental than the learner may realize.

This study has a between participants design. We chose this design because experiments with this design require fewer target items than experiments of the same statistical power with a within participants design. It appeared difficult to find stretches of speech in casual conversations that nearly only contain words that learners know and that are produced at a speech rate appropriate for language learners. Furthermore, it may be more difficult to find language learners willing to participate in a longer experiment, and these participants may not perform optimally at the end of a longer experiment because of tiredness. We therefore opted for a between participants experimental design even though it implies the risk that the effect of the variable manipulated (reduction, in our case) is confounded with coincidental differences between the participants groups.

One of these coincidental differences may be the words that the participants in the two groups knew. If a participant happened not to know all words in a target item’s context, interpretation of the semantic and syntactic cues in the context was more difficult. For some of the speech stretches that we presented, we knew that these contained words that the learners probably did not know. We explained these words to

the learners, but some of the learners may nevertheless have had difficulties understanding them.

As explained above, we have good reasons to assume that the two participant groups had very similar average Dutch proficiencies, including very similar vocabularies. We cannot completely rule out, however, that a small difference in proficiency was present. While all experiments should be replicated, this is especially advisable for experiments comparing groups of language learners. Future research should show whether new experiments, with different participants, different materials, and with different designs produce results confirming what is strongly suggested by this study, which is that second language learners of a high proficiency level also have difficulties processing reduced word pronunciation variants.

Future research could also address another aspect of our experimental design. We allowed participants to listen to a speech stretch more often than once. We hoped that this would elicit participants' best possible performance so that we could see how well they recognize reduced words under optimal conditions. Moreover, by allowing participants to listen to the speech stretches more than once, we minimized the effect of participants' working memory spans, which may differ among participants. Finally, we obtained an additional dependent variable that could show the effect of reduction on speech processing (but which did not for the learners). This feature of our experimental design makes that the results of our experiment are not directly comparable to those from studies where all participants hear the speech stretches equally often (e.g., once). It can also decrease the similarity of the two participant groups. Furthermore, it decreases the ecological validity of our study because in real life listeners cannot hear exactly the same speech stretch several times.

Finally, future research should investigate learners' ability to process reduced words with tasks very different from transcription tasks. Every type of experimental task has advantages and disadvantage and full insight in listeners' processing can therefore only be obtained with the combination of different types of experiments. One of the disadvantages of dictation tasks is that it is not always obvious from the participant's transcription what the listener has understood. Furthermore, it is difficult to investigate with a dictation task what makes the processing of reduced words so hard for learners. Our study should therefore also be complemented with studies testing, for instance, the role of semantic cues with the visual world paradigm.

To wrap up, our results indicate that, like the beginning second language learners tested by Ten Bosch et al. (2016) and like the foreign language learners tested in several other studies (e.g., Nouveau, 2012; Wong et al., 2015), advanced second language learners experience problems recognizing reduced words, which can lead to serious communication problems. If this result is replicated with different participants, different materials, and in different types of experiments, we have to conclude that, in order to learn to process reduced word pronunciation variants, it is not enough for the learner to follow regular language classes and to live in the country where the language is spoken. Learners either need to obtain more experience with casual speech than advanced learners typically have, and / or they need to receive explicit instruction about reduced word pronunciation variants.

4. Acknowledgments

This study received funding from an NWO vici grant (277-70-010) awarded to the first author. The authors would also like to thank the speaker in the dictation task.

References

- Adda-Decker, M., Boula de Mareüil, P., Adda, G., & Lamel L. (2005). Investigating syllabic structures and their variation in spontaneous French. *Speech Communication* 46, 119- 139.
- Brown, J. D., & Hilferty, A. (1986). The effectiveness of teaching reduced forms of listening comprehension. *RELC Journal*, 17, 59-70.
- Council of Europe (2011). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe.
- Dilley, L. C., & Pitt, M. (2010). Altering context speech rate can cause words to appear or disappear. *Psychological Science*, 21, 1664-1670.
- Ernestus, M. (2000). *Voice assimilation and segment reduction in casual Dutch, a corpus based study of the phonology-phonetics interface*. Utrecht: LOT.
- Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua* 142, 27-41.
- Drijvers, L., Mulder, K., & Ernestus, M. (2016). Alpha and gamma band oscillations index differential processing of acoustically reduced and full forms. *Brain and language*, 153, 27-37.
- Ernestus, M. & Warner, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, 39, 253-260.
- Johnson, K. (2004). Massive reduction in conversational American English. Spontaneous speech: data and analysis. *Proceedings of the 1st session of the 10th international symposium* (pp. 29–54). Tokyo: The National International Institute for Japanese Language.

- Mitterer, H., & Tuinman, A. (2012). The role of native-language knowledge in the perception of casual speech in a second language. *Frontiers in psychology*, 3, 249.
- Nouveau, D. (2012). Limites perceptives de l'e caduc chez des apprenants néerlandophones. *Revue Canadienne de Linguistique Appliquée*, 15, 60-78.
- Ranbom, L. J., & Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *Journal of Memory and Language*, 57, 273-298.
- Schuppler, B., Ernestus, M., Scharenborg, O. & Boves, L. (2011). Acoustic reduction in conversational Dutch: A quantitative analysis based on automatically generated segmental transcriptions. *Journal of Phonetics*, 39, 96-109.
- Ten Bosch, L., Giezenaar, G., Boves, L., & Ernestus, M. (2016). Modeling language-learners' errors in understanding casual speech. In G. Adda, V. Barbu Mititelu, J. Mariani, D. Tufiş, & I. Vasilescu (Eds.), *Errors by humans and machines in multimedia, multimodal, multilingual data processing. Proceedings of Errare 2015* (pp. 107-121). Bucharest: Editura Academiei Române.
- Tuinman, A., Mitterer, H., & Cutler, A. (2014). Use of syntax in perceptual compensation for phonological reduction. *Language and speech*, 57, 68-85.
- Van de Ven, M., Ernestus, M., & Schreuder, R. (2012). Predicting acoustically reduced words in spontaneous speech: The role of semantic/syntactic and acoustic cues in context. *Laboratory Phonology*, 3, 455-481.
- Van de Ven, M., Tucker, B.V. & Ernestus, M. (2010). Semantic facilitation in bilingual everyday speech comprehension. In *Proceedings of the 11th Annual Conference*

- of the International Speech Communication Association (Interspeech 2010)*, (pp. 1245-1248). Makuhari, Japan
- Van de Ven, M., Tucker, B.V., & Ernestus (2011). Semantic context effects in the comprehension of reduced pronunciation variants. *Memory and Cognition*, 39, 1301-1316.
- Viebahn, M., Ernestus, M., & McQueen, J. M. (2015). Syntactic predictability in the recognition of carefully and casually produced speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41, 1684-1702.
- Wong, S. W., Mok, P. P., Chung, K. K. H., Leung, V. W., Bishop, D. V., & Chow, B. W. Y. (2015). Perception of native English reduced forms in Chinese learners: Its role in listening comprehension and its phonological correlates. *TESOL Quarterly*.
- Zimmerer, F., & Reetz, H. (2014). Do listeners recover “deleted” final /t/ in German? *Frontiers in psychology*, 5.

Appendix

Orthographic transcription of the dictation task. The words that did not appear in the transcriptions provided to the participants are surrounded by square brackets.

Sentence group 1:

Jij komt niet voor tandartsbezoek daar, toevallig in Alphen? Omdat <beep replacing first name>, die heeft ook uh [dezelfde tandarts als ik].

Sentence group 2:

Ik heb haar ontmoet, dat was natuurlijk op het uh Loufonteins congres he, [met iemand anders]. Daar hield ze ook een lezing, maar ik weet niet [precies] meer waar dat uh over ging.

Sentence group 3:

Ik hoor Frank altijd uh [zaterdagochtend op de radio] maar [ik weet niet of] jij dan wel eens luistert naar het Tros-programma, na Martin Ros komt hij. En dat zet je op als je, de wekker springt uh aan en [dan hoor je die stemmen weer uh].

Sentence group 4:

Maar die uh die praat dan altijd uh heel vrolijk over uh mooie boeken, [alles is prachtig] en mooi enzovoorts. Ook boeken bij zijn eigen, bij zijn eigen uitgeverij die worden [natuurlijk] aangeprezen.

Sentence group 5:

Je kan combi-abonnementen krijgen bij Wolters en Tijdschrift voor Taalbeheersing blijft [gewoon] verschijnen zoals dat uh verschijnt.

Sentence group 6:

Dus je hebt [eigenlijk uh als het goed is] vier tijdschriften dan, ja.

Sentence group 7:

Met die kleine oplage dan uh weet je zeker dat je, de eerste vijftig, nou [die verkopen ze wel]. Brill in Leiden rekent ook uit [hoeveel bibliotheken in de wereld] een boek uh aanschaffen.

Sentence group 8:

En dan uh [maken ze een] calculatie zodanig dat ze quitte spelen met [bibliotheekprijzen] over de wereld en dat uh daar draaien ze nog steeds op.

Sentence group 9:

Nou die rugzakken die uh kosten mij [honderd] maar ik uh als je nou zegt dat je het voor december voor [eind] December bestellen [dan kan ik er wel] uh twintig procent afkrijgen denk ik.

Sentence group 10:

Maar uh [dat moet ik eens] informeren, [ik denk dat dat een] uh mogelijkheid is.

Sentence group 11:

Nou dat was [voor een deel] uh, hij had toch wel uh [verkeerd tentje] of iets dergelijks dat uh ...

Sentence group 12:

Maar hij heeft die uh nacht wel uh overleefd door in die slaapzak te kruipen. Dat is ook een [hele goede slaapzak] uh die dan toch op die hoogte [lekker warm] blijft.

Sentence group 13:

[Jaren geleden] [waren we eens een keer] in België toen uh draaide ook uh de [elektriciteit] op zo'n atoomcentrale. Maar er is geloof ik [vijftig procent] van de energie uh toch vanwege uh atoom hè.

Sentence group 14:

Maar dat [realiseer] je je normaliter niet dat het een uh zo'n groot gedeelte van de [elektriciteit] wordt opgewekt door uh atoomcentrales hè. In [Nederland] is men dat zich [absoluut niet bewust]

Missende woorden

Maar die uh die praat dan altijd uh heel vrolijk over uh mooie boeken, ____ [1] ____ en mooi enzovoorts.
Ook boeken bij zijn eigen, bij zijn eigen uitgeverij die worden ____ [2] ____ aangeprezen.

Slide: 8 / 24

Figure 1: Example of a trial

Missende woorden

Het volgende fragment gaat over een radioprogramma over boeken.

Uitleg van woorden:

Aangeprezen < aanprijzen: zeggen dat iets/iemand erg goed is.
De uitgeverij: een bedrijf dat boeken en tijdschriften publiceert.

Figure 2: Example of the introduction preceding a pair of trials