Dutch vowel production by Spanish learners: duration and spectral features

Pepi Burgos [1], *Mátyás Jani*[2], *Catia Cucchiarini*[1], *Roeland van Hout*[1], *Helmer Strik*[1]

[1] Department of Linguistics, Radboud University Nijmegen, The Netherlands

[2] Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, Hungary

{j.burgos, c.cucchiarini, r.vanhout, h.strik}@let.ru.nl; jani.matyas@itk.ppke.hu

Abstract

In this paper we present a study on Dutch vowel production by Spanish learners that was carried out within the framework of our research on Computer Assisted Pronunciation Training (CAPT). The aim of this study was to obtain detailed information on production of Dutch vowels by Spanish learners, which can be employed to develop effective CAPT programs for this specific target group. We collected speech from learners with varying proficiency levels (A1 - B2 of the CEFR), which was transcribed, segmented and acoustically analyzed. We present data on the frequency of pronunciation errors and on detailed analyses of duration and acoustic properties of the vocalic realizations. The results indicate that Spanish learners of Dutch have difficulties in realizing several Dutch vowel contrasts and that they differ from native speakers in the way they employ duration and spectral properties to realize these contrasts. We discuss these results in relation to those of previous studies on Dutch vowel perception by Spanish listeners and relate them to current theories on speech learning.

Index Terms: L2 phonology acquisition, language learning, Computer Assisted Pronunciation Training (CAPT)

10.21437/Interspeech.2014-130

1. Introduction

Research on second language (L2) learning has shown that adult learners seldom achieve native speaker pronunciation [1], [2]. Difficulties in acquiring L2 phonology appear to be related to problems in perception, production or both [3], [4], [5], [6], which are often caused by interference from the native language (L1). Practicing pronunciation is often neglected in language lessons because it is too time-consuming. For this reason, there is growing interest in developing advanced CAPT systems in which Automatic Speech Recognition (ASR) is used to provide practice and feedback in L2 pronunciation in an effective and efficient way.

In our research we investigate to what extent a CAPT system that is designed to address L1-L2 pronunciation problems can help L2 learners improve their pronunciation. The languages in question are L1 Spanish and L2 Dutch, which is an interesting combination because of the phonological differences (see Section 2). To be able to develop such a Spanish-Dutch specific system, we need information about which aspects of Dutch phonology are problematic for Spanish learners. In Burgos et al. [7], [8] we reported on a first investigation of Spanish accented Dutch.

In this paper we report on a follow-up study in which Dutch L2 vowel pronunciation by Spanish speakers is explored in more detail. In particular, the aim of the present study is to investigate vowel productions by Spanish learners of Dutch with a view to determining which vowel contrasts are more problematic and which dimensions underlie the difficulties in production.

This paper is organized as follows. In Section 2 we describe the background to our research. Section 3 gives details on the method, corpus material and acoustic analysis. Section 4 is devoted to the results which are discussed in Section 5. Finally, in Section 6 conclusions are drawn.

2. Research background

The Dutch and Spanish vowel systems differ in important respects [7], [8], [9], [10]. First, Spanish has five vowels (/a, e, i, o, u/) [11], whereas Dutch has fifteen unreduced vowels (seven tense vowels: /i, y, u, e, ø, o, a/; five lax vowels: /I, ɛ, ɔ, ʏ, ɑ/; three diphthongs: /ɛi, oey, ɔu/) and the reduced vowel schwa /ə/ [12]. Second, Dutch has a tense/lax distinction including vowel length, whereas Spanish does not have contrastive vowel length. Third, Dutch has four front rounded vowels: /ʏ, y, ø, oey/, whereas in Spanish all rounded vowels (/o, u/) are back.

Most studies on Spanish native listeners' difficulties with Dutch vowels concern speech perception [9], [10], [13], [14], while limited research has been conducted on vowel production. Previous research at our lab [7], [8] has shown that Spanish learners have difficulties in producing Dutch vowels and that L1 phonology influences L2 vowel production as the five Spanish vowels function as attractors.

To develop effective CAPT systems it is important to get a better understanding of the pronunciation problems observed. Research on Dutch vowel perception has shown that native speakers of Spanish employ duration and spectral cues in distinguishing Dutch vowels differently from Dutch native speakers [9], [13], [15]. It is therefore interesting to find out how Spanish learners of Dutch employ spectral features and duration in producing Dutch vowel contrasts. To this end, the follow-up study was conducted which is reported on in the remainder of this paper.

3. Method

3.1. Speech material

For our research we decided to collect a more focused corpus of Spanish L1 Dutch L2 than the corpus of extemporaneous in [7], [8]. This new corpus should contain read speech and sufficient productions of Dutch speech sounds that are problematic for Spanish learners. For the present study, we used a subset of this corpus in which Spanish learners of Dutch read monosyllabic words from a computer screen. The elicitation material was the same as that used in a study by Van der Harst [16], consisting of a set of twenty-nine monosyllabic Dutch words. The words contained all fifteen Dutch monophthongs in stressed position followed either by /s/ or /t/, as it is known that alveolar consonants minimally change the quality of the preceding vowels [16], [17].

3.2. Participants

Twenty-eight adult Spanish learners of Dutch from Spain and various Latin American countries, nine males and nineteen females, took part in this study. They had been living in The Netherlands for an average of 4.82 years (ranging from one month to twenty years) at the time of this study. All participants had taken Dutch courses at some point in time during their stay in The Netherlands. Some of them were participating in Dutch courses at Radboud in'to Languages, the language learning centre of the Radboud University [http://www.ru.nl/radboudintolanguages] when recordings for this study took place. Others were not participating in Dutch courses at the time of the recordings, however they reported being exposed to Dutch and using it daily. All participants were familiar with the CEFR (Common European Framework of Reference) [18] and rated their proficiency level of Dutch (A1, n=10; A2, n=7; B1, n=4; B2, n=7) according to the CEFR Self-Assessment Grid [19].

3.3. Analyses of the speech recordings

The words read by the speakers were transcribed in Praat [20]. SAMPA (Speech Assessment Methods Phonetic Alphabet) [21] phonetic alphabet was used for all transcriptions. Pronunciation variants of the lexicon of the Spoken Dutch Corpus [22] were used to generate an automatic phonemic transcription that contained the canonical transcription of the words in question. The phonemes that were actually realized were indicated in a different tier where transcribers rated every phoneme and annotated deletions, substitutions and insertions. After the annotation was completed we brought in another transcriber (a native speaker of Dutch) who was asked to check the accuracy of a subset (10% of all recordings) of the annotated material by the first transcriber. Subsequently, the intertranscriber agreement was calculated by comparing 10% of the material that had been transcribed by both annotators. The degree of agreement between the first and the second annotator was acceptable, as kappa was .785. The annotations were used to calculate confusion matrices and obtain information about the frequency of pronunciation errors made by Spanish learners.

We then proceeded to vowel segmentation following the procedures described in [23] and [16]. In segmenting vowels, we looked at information from the waveform, spectrogram, formant tracks and auditory cues to determine the beginning and the end of each vowel. The segmentation of vowels was then checked by a Dutch native phonetician. Since there were some discrepancies, the first transcriber and the Dutch native phonetician went through 50% of the material each and, where necessary, altered the beginning and the end of vowels that had already been segmented.

3.4. Acoustic analyses

To get a better understanding of the nature of the pronunciation errors we proceeded to feature extraction for every vowel. Duration, pitch (F0), F1 and F2 were extracted and analyzed. Acoustic analyses were carried out automatically to obtain measurements of formants, duration and pitch. All measurements automatically extracted were manually checked and, where errors were found, corrected. Subsequently, we normalized all vowel realizations by using the Lobanov transformation in order to neutralize the formant frequency variations resulting from anatomic differences among informants [24]. The normalized values for the Dutch vowels were taken from Van der Harst [16]. For the Spanish learners normalizations were based on the vowels that were realized as /i/, /u/, /ɛ/, /ɔ/, and /a/, as these vowels reflect both the Spanish vowel system and represent the periphery of the vowel space. For duration we used the raw values.

4. Results

4.1. Error frequency

Table 1 displays information on the mispronunciations of the twelve Dutch monophthong vowels. The percentage of errors is given plus the most frequent mispronunciation and its percentage of occurrence.

Table 1. *Frequency of vowel mispronunciations and most frequent error per target vowel; Vow =Target vowel.*

Vow	Example	%Error	Most Frequent Error	%Most Frequent Error
/ɑ/	*Rat*	71.40%	/a/	71.40%
/ø/	*Neus*	58.90%	/əʏ/	30.40%
/y/	*Fuut*	50.00%	/u:/	28.60%
/ɪ/	*Vis*	46.40%	/i/	46.40%
/o/	*Boot*	44.60%	/ɔ:/	28.60%
/e/	*Mees*	39.30%	/ɛ:/	21.40%
/u/	*Voet*	32.10%	/u:/	10.70%
/ʏ/	*Bus*	25.00%	/u/	19.60%
/a/	*Staat*	25.00%	/a:/	19.60%
/ɔ/	*Vos*	10.70%	/ɔ:/	5.40%
/ɛ/	*Vet*	10.70%	/ɛ:/	5.40%
/i/	*Kies*	7.10%	/e/	1.80%

From this table it is clear that an appreciable number of mispronunciations are made by Spanish learners. Overall these percentages can be qualified as rather high. If we look at the column with the most frequent mispronunciations, we see that, with the exception of /əʏ/, vowels close to the five Spanish vowels figure prominently in this list. The /əʏ/ mispronunciations can probably be ascribed to influence from L1 orthography as the Dutch phoneme /ø/ is represented by two graphemes <eu>. In Spanish there is a one-to-one correspondence between graphemes and phonemes [10], which can explain why two graphemes are also realized as two phonemes. In Burgos et al. [7], [8] we already discussed the role of orthography, as part of the substitutions we observed could be explained in this way.

The prominent position of vowels close to the five Spanish vowels in the list of mispronunciations lends support to the idea of these vowels functioning as "attractors", as advanced in [7], [8]. This is also in line with results of experiments on Dutch vowel perception by Spanish listeners [13].

The important question now is how these vowels fulfil their role of attractors. Again, in many cases it can be just a question of orthographic interference. For instance, we can see in Table 1 that /u/ and /u:/ are often realized instead of /y/ or /ʏ/, which are clearly instances of so-called spelling pronunciation, as in Spanish the grapheme <u> stands for the vowel /u/.

Orthography might also be responsible for the high number of long vowels, probably triggered by the frequent, (syllable-dependent) presence of two graphemes in Dutch for the vowels /a/, /e/, /o/, /y/ and /u/, as <aa>, <ee>,<oo>, <uu> and <oe>, respectively [10].

Table 1 shows that Spanish learners do employ duration as a dimension to realize contrasts between /ɑ/ - /a/, /I/ - /e/, /ɔ/ - /o/ and /ʏ/ - /ø/ in Dutch, even though this dimension is not used distinctively in their L1. This is in line with the results of Dutch vowel contrast perception experiments with Spanish listeners [9]. An interesting question at this point seems to be how Spanish learners exactly use duration and spectral features to realize Dutch vowel contrasts. Symbolic representations like the annotations contained in Table 1 have limited capabilities in this respect: the transcriber chooses the symbol that seems more appropriate and this might correspond to a realization in between two Dutch phonemes which cannot accurately be indicated by the symbol in the set available.

To get a better understanding of Dutch vowel production by Spanish learners, we decided to analyse three vocalic contrasts in more detail in this study, i.e. the contrasts /ɑ/ - /a/, /I/ - /i/ and /ʏ/ - /ø/, for which the dimensions of duration and spectral features are used differently by Dutch native speakers (see Table 2).

Table 2. *Vowel contrast and their distinctive features, as employed by Dutch native speakers; + / - distinction.*

Vowel pair	Duration	Place (F1/F2)
/ɑ/ - /a/	+	+
/I/ - /i/	-	+
/ʏ/ - /ø/	+	-

As is clear from this table, the vowels in the first pair, /ɑ/ - /a/, differ from each other in the two dimensions place and duration, while the /I/ - /i/ contrast hinges on place and /ʏ/ - /ø/ on duration [15]. In the following section we investigate how these contrasts are realized by Spanish learners.

4.2. Duration and spectrum of three vowel contrasts

The duration, F1 and F2 values of the three vowel pairs were calculated per vowel. The vowels in each pair differ significantly in duration. The mean and standard deviation (SD) for the vowel /ɑ/ were 154.60 ms, SD=27.99, and for the /a/ 263.44 ms, SD=59.19. The ANOVA gave a significant result with an eta squared (i.e. explained variance: minimum 0, maximum 1) value of .754, indicating a strong effect. The means and SD for the vowel /I/ were 108.91 ms, SD=22.76 and for the /i/ 151.42 ms, SD=40.40. This difference is significant with an eta squared value of .582, a medium sized effect indicating that many Spanish learners make a duration difference that is not present in native Dutch. The mean and SD for the vowel /ʏ/ were 124.96 ms, SD=34.72 and for /ø/ 231.67 ms, SD=38.49. The difference is statistically significant, the eta squared being extremely high, i.e. .906. Almost all Spanish learners produced a duration difference. The conclusion is that the Spanish learners were successful in picking up the native Dutch difference in vowel duration for the /ɑ/ - /a/ pair and the /ʏ/ - /ø/ pair, but they erroneously tend to apply the length distinction to the /I/ -/ i/ pair.

Which F1 and F2 values do our learners realize? These are shown in the plot of Figure 1, including the Dutch native target values based on [16]. Two vowel pairs seem to be largely merged. The /I/ - /i/ vowels in the upper left corner of Figure 1 overlap for the most part, the /I/ being mostly realized

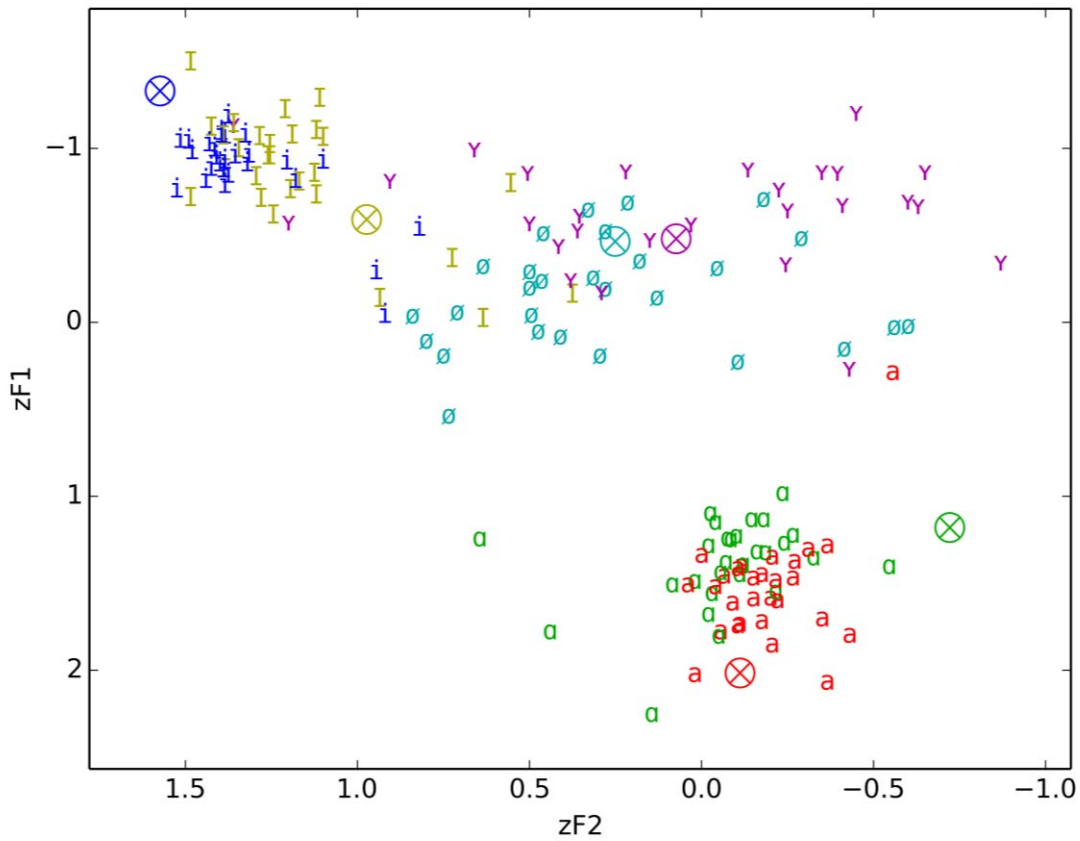

Figure 1. *Scatterplot of vowel realizations by Spanish learners; the Dutch values (crossed circles) from Van der Harst [16].*

in the /i/ area, outside the Dutch target value. A few /I/ realizations are lower, but the same applies to a few /i/ realizations. The F1 and F2 differences were tested by ANOVA. Only the F2 value gave a statistically significant result with a low eta squared value of .174.

The /ɑ/ - /a/ pair in the lower right corner of Figure 1 shows that the Dutch target of /ɑ/ remains out of scope. The realizations are more similar to the /a/ target, albeit somewhat higher. The F1 values are more in line with /ɑ/, the F2 values with /a/. It is remarkable that several /ɑ/ realizations are too fronted, with too high F2 values. ANOVA tests of the F1 and F2 differences show statistically significant results for F2 with a low eta squared value of .166, but this result is the opposite of the effect expected, as the extreme fronted /ɑ/ realizations make the average F2 value higher than that for /a/.

The third pair, /ʏ/ - /ø/, in the upper middle part of Figure 1, gives a completely different picture, with values scattered over the whole F2 range. There is variation on the F1 dimension too, but it is much less extreme. The F2 dispersion means that the roundedness of these two vowels is not well captured by the Spanish learners. Only the F1 value gave a statistically significant result with a low eta squared value of .174. The target Dutch vowel pair has no such distinction, meaning that the F1 values of the /ʏ/ are too high, with the result that several realizations are near the area of /u/.

Table 3 shows the distinctions the Spanish learners make. When we compare Table 2 (Section 4.1) with Table 3 we can see that Spanish native speakers do not employ front rounding, durational and spectral cues in the same way as native speakers do.

Table 3. *Vowel contrasts and distinctive features, as employed by Spanish learners; wrong contrasts in blue; V pair = Vowel pair, Dur = Duration.*

V pair	Dur	Place (F1/F2)
/ɑ/ - /a/	+	- (F1) , + (F2) but opposite direction
/I/ - /i/	+	- (F1) , + (F2)
/ʏ/ - /ø/	+	+ (F1) , - (F2) but overdispersion

5. Discussion

The acoustic analyses of Spanish learners' realizations of the Dutch vowel contrasts: /ɑ/ - /a/, /I/ - /i/ and /ʏ/ - /ø/ presented in the previous sections convincingly show that Spanish learners have difficulties in realizing Dutch vowel contrasts. The results indicate that Spanish learners tend to rely more on duration (see Table 3, duration) than on spectral features (see Table 3, F1 and F2 values) to distinguish vowel in a pair. They even use duration distinctions where native do not employ them (the case of /I/ - /i/). The duration distinction in the /ɑ/ - /a/ and /ʏ/ - /ø/ match native pronunciation in reading words.

Concerning spectral properties, the Spanish learners' spectral distinctions between the vowels involved in these pairs generally have wrong values, different from those in native speech. The /ɑ/ - /a/ shows an F2 distinction only, but in the opposite direction. The /I/ - /i/ gives only a weak, gradual distinction on the F2, whereas Dutch native speakers use the frequency spectrum systematically to distinguish the two vowels (see also [13], [14]). The /ʏ/ - /ø/ should not produce a spectral difference, but gives a weak, gradual F1 effect, next to an overdispersed distribution of F2 values, over the whole front to back range.

The realizations of the vowel pairs we have analyzed have either too low spectral contrastive values in relation to the five native Spanish vowel categories (/a, e, i, o, u/), as we have seen in the contrasts /ɑ/ - /a/ and, /I/ - /i/, or they have values in between, in the case of the rounded front vowels /ʏ/ - /ø/, with the consequence that the realizations cover the full F2 range and are overdispersed. While in the case of /ɑ/ - /a/ and /I/ - /i/, there is a specific Spanish vowel, the phoneme /a/ and the phoneme /i/, respectively, on which both Dutch vowels can be mapped, for the Dutch vowels /ʏ/ and /ø/ there is no direct Spanish counterpart. As Goudbeek et al. [15, page 123] put it: /ø/ and /ʏ/ "are unfamiliar and fall in an empty portion of the native vowel space."

A plausible explanation for the different treatment of the vowel pairs /ɑ/ - /a/, /I/ - /i/ and /ʏ/ - /ø/ by Spanish learners appears to be related to how L2 sound contrasts are mapped onto native categories, as explained in the SLM (Speech Learning Model) by Flege [3], the PAM (Perceptual Assimilation Model) by Best [6], [25] and L2LP (Second Language Linguistic Perception Model) by Escudero [26]. According to these models, L2 vowel contrasts that are mapped onto one single category, as in the case of /ɑ/ - /a/ and, /I/ - /i/, are expected to be more difficult than contrasts that are mapped onto separate Spanish vowel categories (see also [13]). The mispronunciation frequency data in our study lend support to this interpretation. However, considering that these frequencies are all rather high, it seems better not to draw more definite conclusions on the relative difficulty in learning and producing the distinctions we investigated before we have included and analyzed production data on the other Dutch vowels.

6. Conclusions

The study we conducted was aimed at obtaining detailed information about Dutch vowel production by Spanish learners of Dutch. Based on analyses of mispronunciation frequency and on durational and spectral features of the vocalic realizations we can draw the following conclusions. First, the influence of the native phonology on vowel production is pervasive. Specifically, the results seem to confirm the role of the restricted set of Spanish vowels as attractors for the larger set of Dutch vowels. Second, when realizing Dutch vowels Spanish learners of Dutch do not use durational and spectral features in the same way as Dutch native speakers do. Third, in producing vowels that differ both in duration and spectral features, Spanish learners tend to rely more on duration to realize the contrast.

7. Acknowledgements

This research was supported by the Centre for Language Studies (CLS) at the Radboud University Nijmegen, The Netherlands. We would like to thank Joop Kerkhoff for his help in the analysis of the speech files, and Sander van der Harst for making the acoustic values of the Dutch speakers available to us.

8. References

[1] Long, M., “Maturational constraints on language development", Studies in Second Language Acquisition, 12:251-285, 1990.

[2] Birdsong, D. and Molis, M., “On the evidence for maturational constraints in second language acquisition”, Journal of Memory and Language, 44(2):235-249, 2011.

[3] Flege, J.E. , “Second language speech learning: Theory, findings and problems”, in W. Strange [Ed], Speech perception and linguistic experience: Issues in cross-language research, 233-227, York Press, 1995.

[4] Flege, J., “A critical period for learning to pronounce foreign languages?”, Appl. Ling. 8:162-177, 1987.

[5] Flege, J., “Age of learning and-second-language speech”, in D. Birdsong [Ed], Second Language Acquisition and the Critical Period Hypothesis, 101-132, Lawrence Erlbaum Associates, 1999.

[6] Best, C.T., “A direct realist view of speech cross language speech perception”, in W. Strange [Ed], Speech perception and linguistic experience: Issues in cross- language research, 71-206, York Press, 1995.

[7] Burgos, P., Cucchiarini, C., Van Hout, R. and Strik, H., “Pronunciation errors by Spanish learners of Dutch: A data-driven study for ASR- based pronunciation training”, Proceedings of Interspeech 2013, Lyon, France, 2385-2389, 2013.

[8] Burgos, P., Cucchiarini, C., Van Hout, R. and Strik, H., “Phonology acquisition in Spanish learners of Dutch: Error patterns in pronunciation”, Language Sciences, 41:129-142, 2014.

[9] Escudero, P., Benders, T. and Lipski, S., "Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners", Journal of Phonetics 37:452-465, 2009.

[10] Escudero, P. and Wanrooij, K., “The effect of L1 orthography on non-native vowel perception”, Language and Speech, 53(3):343-365, 2010.

[11] Hualde, J.I., “The sounds of Spanish”, Cambridge University Press, 2005.

[12] Booij, G., “The phonology of Dutch”, Clarendon Press, 1995.

[13] Escudero, P., and Williams, D., “Perceptual Assimilation of Dutch vowels by Peruvian Spanish listeners”, Journal of the Acoustical Society of America, 129(1):1-7, 2011.

[14] Escudero, P., and Williams, D., “Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch ”, Journal of the Acoustical Society of America Express Letters, 131(5):406-412, 2012.

[15] Goudbeek, M., Cutler, A. and Smits, R., “Supervised and unsupervised learning of multidimensionally varying non-native speech categories“, Speech Communication 50:109–125, 2008.

[16] Van der Harst, S., “The vowel space paradox: A sociophonetic study on Dutch”, Ph.D. Thesis, LOT, 2011.

[17] Van Hout, R., De Schutter, G., De Crom, E., Huinck, W., Kloots, H. and Van de Velde, H., “De uitspraak van het Standaard-Nederlands. Variatie en varianten in Vlaanderen en Nederland”, in E. Huls, B. Weltens [Eds], Artikelen van de Derde Sociolinguïstische Conferentie, 183-196, Eburon, 1999.

[18] CEFR (Common European Language of Reference for Languages). Online: http://www.coe.int/t/dg4/linguistic/cadre1_en.asp, accessed on 4 Mar 2014.

[19] CEFR (Common European Language of Reference for Languages),Self-Assesment Grid. Online: http://www.coe.int/t/dg4/education/elp/elp-reg/Source/assessement_grid/assessment_grid_ES.pdf, accessed on 4 Mar 2014.

[20] Boersma, P. and Weenink, D., “Praat: Doing phonetics by computer”, http://www.praat.org., 2010.

[21] Wells, J., “SAMPA. Computer Readable Phonetic Alphabet”, http://www.phon.ucl.ac.uk/home/sampa, 2004.

[22] Oostdijk, H.J., “The design of the Spoken Dutch Corpus”, in P. Peters, P. Collins, A. Smith, [Eds], New Frontiers of Corpus Research, Rodopi, 105-112, 2002.

[23] Van Son, R.J., Binnenpoorte, D., Van den Heuvel, H. and Pols, L.C., “The IFA corpus: a phonemically segmented Dutch open source speech database”, Proceedings of Eurospeech 2001, Aalborg, Denmark, (3):2051-2054, 2001.

[24] Adank, P., Smits R. and Van Hout, R., “A comparison of vowel normalization procedures for language variation research”, Journal of the Acoustical Society of America 116(5): 3099-3107, 2004.

[25] Best, C. T. and Tyler, M.D., “Non-native and second-language speech perception: Commonalities and complementarities”, in M.J. Munro, O.-S. Bohn [Eds], Second Language Speech Learning: The role of language experience in speech perception and production, 13-34, John Benjamins, 2007.

[26] Escudero, P., “Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization”, Ph.D. thesis, LOT, 2005.