

Non-native pronunciation:

Patterns of learner variation in Spanish-accented Dutch

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111

e-mail: lot@uu.nl
<http://www.lotschool.nl>

Cover illustration: Detail of 'Drie Levensbomen' by Lieske Stieger-Brouwers. All rights reserved.

ISBN: 978-94-6093-268-7
NUR 616

Copyright © 2018: Pepi Burgos. All rights reserved.

Non-native pronunciation:

Patterns of learner variation in Spanish-accented Dutch

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,

volgens besluit van het college van decanen

in het openbaar te verdedigen op vrijdag 26 januari 2018

om 11.30 uur precies

door

Josefa María Burgos y Guillén

geboren op 26 december 1970

te Málaga (Spanje)

Promotor:

Prof. dr. R.W.N.M. van Hout

Copromotor:

Dr. B.C. Planken

Manuscriptcommissie:

Prof. dr. M.T.C. Ernestus

Prof. dr. H.M.G.M. Jacobs

Prof. dr. H. Quené (Universiteit Utrecht)

dr. J. Cebrian (Universitat Autònoma de Barcelona, Spanje)

dr. J.C. Mora (Universitat de Barcelona, Spanje)

A José y Francisca,
las estrellas que siempre me acompañan

Contents

Contents	vii
1 Introduction	1
1.1 Introduction	1
1.2 Background	2
1.3 Spanish and Dutch	4
1.4 The phonology of Spanish and Dutch	8
1.4.1 Vowels	8
1.4.2 Consonants	12
1.5 Pronunciation problems of Spanish learners of Dutch, English, German and French	14
1.6 Speech perception models and Dutch L2 perception of Spanish learners	19
1.7 Perception of foreign-accented speech by native listeners	27
1.8 Research objective and design	29
1.9 Outline of the dissertation	32
2 Phonology acquisition in Spanish learners of Dutch: Error patterns in pronunciation	37
2.1 Introduction	37
2.2 Research background	39
2.2.1 L2 speech perception models	39
2.2.2 Orthographic interference	39
2.2.3 The phonology of Spanish and Dutch	40
2.2.3.1 Vowels	40
2.2.3.2 Consonants	44
2.2.4 Production and perception difficulties of Spanish learners of Dutch	44
2.2.5 Expected difficulties	46
2.2.5.1 Vowels	46
2.2.5.2 Consonants	46
2.3 Method	47
2.3.1 Participants	47
2.3.2 Speech material	47
2.3.3 Data analysis	48
2.4 Results	50
2.5 Discussion	58
2.5.1 Vowels	58
2.5.2 Consonants	63
2.6 Conclusions and future perspectives	64
3 Spanish-accented Dutch vowel productions: Duration and spectral features	68
3.1 Introduction	68

3.2	Five predictions	75
3.3	Method	79
3.3.1	Spanish learners	79
3.3.2	Speech material	83
3.3.3	Procedure	85
3.3.4	Analysis of the speech recordings	86
3.3.5	Acoustical analyses	86
3.4	Results	88
3.4.1	Duration	88
3.4.2	Spectral features	91
3.4.2.1	Monophthongs	91
3.4.2.2	Long mid vowels	96
3.4.2.3	Diphthongs	99
3.5	Discussion	102
3.6	Conclusions	106
4	<i>Auris populi</i> : Crowdsourced native transcriptions of Dutch vowels spoken by adult Spanish learners	108
4.1	Introduction	108
4.2	Research background	109
4.3	Method	110
4.3.1	Speakers	111
4.3.2	Speech stimuli	111
4.3.3	Listeners	113
4.3.4	The crowdsourcing experiment	113
4.3.5	Quality control	114
4.4	Results	114
4.4.1	Listeners' transcriptions, vowel confusions	114
4.4.2	Listeners' transcriptions of three vowel pairs	117
4.5	Discussion	117
4.6	Conclusions	120
5	Native listeners' transcriptions of vowel variation in L2 speech	122
5.1	Introduction	122
5.2	Method	129
5.2.1	Listeners	129
5.2.2	Speech stimuli	129
5.2.3	Procedure	132
5.2.4	Intra-transcriber consistency	133
5.3	Results	134
5.3.1	The non-native confusion matrix	134
5.3.2	Comparison crowdsource and present study	137
5.3.3	The native confusion matrix	139
5.4	Discussion	141

5.5	Conclusions	146
6	Matching acoustical properties and native perceptual assessments of L2 speech	148
6.1	Introduction	148
6.2	Method	154
6.2.1	Spanish learners	154
6.2.2	Stimulus materials	156
6.2.3	Analysis of the speech recordings	158
6.2.4	Acoustical analyses	159
6.2.5	Native Dutch listeners	160
6.3	Results	161
6.3.1	Acoustic data	161
6.3.1.1	Non-native data	162
6.3.1.2	Native data	169
6.3.2	Listener data	172
6.3.2.1	Non-native data	173
6.3.2.2	Native data	176
6.3.3	Comparison acoustic and listener data	178
6.3.3.1	Spanish learners as a group	178
6.3.3.2	Individual differences across learners	182
6.3.3.2.1	Acoustic data	182
6.3.3.2.2	Listener data	187
6.3.3.2.3	Comparison acoustic and listener data	192
6.4	Discussion	196
6.5	Conclusions	202
7	Conclusion and discussion	204
7.1	Introduction	204
7.2	Answering the research questions	205
7.3	Individual variability	210
7.4	Adaptive mechanisms (statistical classification and human recognition)	215
7.5	L2 language proficiency and pronunciation	218
7.6	Evaluating the speech perception models	221
7.7	Limitations	224
7.8	Future prospects	225
7.9	Societal relevance	229
	References	230
	Samenvatting (Dutch summary)	242
	Resumen (Spanish summary)	253
	Acknowledgements	267
	Publications	269
	Curriculum vitae	271

1 Introduction

1.1 Introduction

Learning Dutch is hard for adult Spaniards, and when asked what the most difficult aspect of learning Dutch is, most of them would probably answer: “*la pronunciación*”, ‘the pronunciation’. Indeed, native Dutch listeners often seem to struggle to understand Dutch words uttered by Spanish learners. As a result, my Dutch friends seem to think that my Spanish friends do not put enough effort into pronouncing Dutch accurately, whereas my Spanish friends seem to think that my Dutch friends do not put enough effort into trying to understand what they intend to say. Watching such uneasy interactions between my Dutch and Spanish friends motivated me to investigate the pronunciation problems of adult Spanish learners of Dutch, and their possible sources, as well as to find out how well native Dutch listeners perceive Spanish-accented Dutch pronunciation.

This chapter provides an introduction to the present investigation and contextualizes it within the literature. It starts with information on adult Spanish migrant workers in the Netherlands, in the past and more recently. It then describes the characteristics of Spanish and Dutch and the phonology of Spanish and Dutch, both for vowels and consonants. Subsequently, the pronunciation problems of Spanish learners of Dutch, English, German and French are discussed. After a consideration of current speech perception models and how these can be used to understand Dutch L2 perception by Spanish learners, the objectives and design of the present investigation are introduced. The chapter ends with an outline of the remaining chapters of this dissertation.

1.2 Background

Globalization has brought about new global cultural forms, media, and communication technologies, which are reshaping interaction within and across cultural and linguistic settings. Globalization and increasing mobility have had a dramatic impact on education and the way people learn languages. The necessity to learn a foreign or a second language (L2) effectively is essential for people who operate in an international context or live in an L2 environment. Adult learners face many challenges in acquiring an L2. These include the acquisition of new morphological paradigms, syntactic structures, lexical items and phonological properties. Although L2 speakers may be fluent, their accented speech is unlikely to be as intelligible¹ as native speech, which could affect the effectiveness of communication (Cutler, 2012; Van Wijngaarden 2001). It is well known that adult L2 learners have great difficulty in mastering L2 speech sounds (Birdsong & Molis, 2011; Long, 1990;), and many of them retain a foreign accent even after having spent several years in the host country². Having a foreign accent can have social repercussions. Adult L2 learners' competences are commonly judged on the basis of their foreign accent, which can be disadvantageous for career opportunities, successful interaction and social acceptance (Brennan & Brennan, 1981; Lippi-Green, 1997; Moyer, 2013).

The Netherlands has seen several waves of worker migration. After the Second World War, the Netherlands was in ruins: its industry had been destroyed and its towns and cities devastated. By the end of the 1950s and early 1960s, post-war reconstruction had led to an acute shortage on the labour

¹ According to Munro and Derwing (1995), intelligibility can be defined as “the extent to which a speaker’s message is actually understood by a listener” (p.76).

² The significance of the variable length of residence (LoR) for L2 foreign accent is rather unstable across studies (see Piske, MacKay, & Flege, 2001).

market. Individual companies as well as the Dutch government began to recruit so-called ‘guest workers’ (in Dutch: *gastarbeiders*) from Southern Europe (especially Spain, Italy and Turkey) and North Africa (Morocco). This first wave of migrants consisted mostly of low-skilled workers. Spanish migrants worked in the metal industry, the blast furnace industry and the Port of Rotterdam. Most found employment at Philips, the multinational Dutch electronics corporation in Eindhoven³. In fact, by 1966, Philips had built two separate community villages (i.e., *El Pinar*, ‘the pine tree’ and *El Prado*, ‘the meadow’) to house its Spanish workers on the outskirts of the city of Eindhoven, as they were not supposed to mingle with the Dutch population. The company also provided them with so-called *centros*, ‘community centres’ (e.g., *centro español*, ‘Spanish community centre’) where guest workers could socialize with each other, watching Spanish movies, reading Spanish newspapers and eating Spanish homemade food. The *centro español* still exists today. By 1974, a total of 32,000 Spaniards had come to the Netherlands. But with the death of the Spanish dictator Francisco Franco in 1975 and the transition to a Spanish democracy, many returned home. This was not the case for Moroccan and Turkish guest workers, who stayed on in the Netherlands, particularly after the introduction of the Dutch law on family reunification in 1974 which enabled the families left behind by guest workers to migrate to the Netherlands as well⁴.

³ More information about the *Colonia española de Eindhoven*, ‘Spanish working community in Eindhoven’ can be found in <http://www.emigracioneindhoven.dse.nl/> (date last viewed 31/08/17).

⁴ See also *Vijf eeuwen migratie*, ‘five centuries of migration’, in <http://www.vijfleeuwenmigratie.nl/term/Gastarbeiders>, for more information about migration in the Netherlands (date last viewed 31/08/17).

The second wave of Spanish migration to the Netherlands began nearly ten years ago, following the economic crisis in 2008 and the steep rise in the Spanish unemployment rate⁵. These Spanish migrants differ from the Spanish guest workers in the 1960s: they are highly skilled, mobile, highly educated, and speak English well. Most of them work in the high-tech and healthcare sectors or have enrolled as students in Dutch higher education⁶. However, most do not have linguistic knowledge of Dutch when they arrive in the Netherlands. While they can get along communicating in English at first, they soon become aware of the importance of speaking Dutch, because it is required at work or because they want to improve their social interaction.

1.3 Spanish and Dutch⁷

Spanish and Dutch are languages which differ from each other in important respects. Spanish belongs to the Romance language family, together with Italian, French and Romanian. Spanish is more similar to Catalan, Galician and Portuguese, as they all share the influence of an Iberian substratum and a Moorish superstratum (Lapesa, 1981). Dutch, on the other hand, belongs to

⁵ According to the Dutch Social and Cultural Planning Office (*Sociaal en Cultureel Planbureau, SCP*), there were approximately 40,000 Spaniards registered as living in the Netherlands in 2015. More information can be found in <https://www.scp.nl/Zoekresultaten?searchbase=0&searchrange=10&searchpage=1&freetext=Summary+New+Spanish+migrants+in+the+Netherlands&submit=Zoeken> (date last viewed 31/08/17).

⁶ 65% of Spanish migrants working in the Netherlands are employed in professional or technical jobs (e.g., as researchers, teachers, engineers, nurses or IT specialists) (*SCP*, 2016).

⁷ In this dissertation the term “Dutch” refers by default to “Northern Standard Dutch”, the language variety spoken in the Netherlands, which is the focus of the present investigation, not to be confused with “Southern Standard Dutch”, i.e., the language variety spoken in Flanders, Belgium.

the Germanic language family⁸, and more specifically to the group of West Germanic languages, together with English and German. We will briefly discuss some differences in morphology, syntax, syllable structure and rhythm between Spanish and Dutch.

Spanish morphology, which differs substantially from Dutch morphology, can be very difficult for L2 learners of Spanish because of its complexity. Affixation, i.e., the process of adding a morpheme – or affix – to a word to create either a different form of that word (e.g., *casa*, ‘house’, *casita* ‘little house’) or a new word with a different meaning (*poner*, ‘to put’, *imponer*, ‘to impose’) is very frequent in Spanish. Spanish has different types of morphemes. For nouns and adjectives, there are gender (masculine and feminine; e.g., *chico*, ‘boy’, *chica*, ‘girl’) and number morphemes (e.g., *curso*, ‘course’, *cursos*, ‘courses’; *color*, ‘colour’, *colores*, ‘colours’), whereas for verbs there are mode, time, voice, aspect, person and number morphemes (e.g., *bailarán*, ‘they will dance’). Dutch morphology is not as rich as Spanish morphology, especially when it comes to verbal affixation. Dutch morphology has morphemes for nouns (e.g., *cursus*, ‘course’, *cursussen*, ‘courses’; *appel*, ‘apple’, *apples* ‘appels’) and verbs (e.g., *ik werk*, ‘I work’, *ze werken*, ‘they work’), as well as morphemes to form diminutives (e.g., *huis*, ‘house’, *huisje*, ‘little house’) or to create a new word with a different meaning (e.g., *leggen*, ‘to put’, *uitleggen*, ‘to explain’). Generally, Spanish learners of Dutch do not have serious difficulties with Dutch morphology, whereas Dutch learners of

⁸ Traditionally, Germanic languages are divided into three groups: West Germanic, including English, German, and Dutch (Northern and Southern Standard Dutch), North Germanic, comprising Danish, Swedish, Icelandic, Norwegian and Faroese, and East Germanic, now extinct, including only Gothic and the languages of the Vandals, Burgundians and a few other tribes (cf. König & Van der Auwera, 1994).

Spanish do have problems with the complex Spanish morphology, especially when it comes to verbal affixation.

Regarding syntax, Spanish syntax is much easier for L2 learners of Spanish, including Dutch learners of Spanish, than Spanish morphology. The default word order pattern in Spanish is [S]VO (Subject⁹, Verb, Object), a pattern which in fact does not change when prepositional phrases of time, manner and place, and other verbs are added to the sentence. For instance, a possible word order could be *Mañana compro el libro en Granada*, ‘Tomorrow I will buy the book in Granada’, although other word orders like *Compro el libro en Granada mañana*, ‘I will buy the book in Granada tomorrow’ are also grammatically correct. In other words, Spanish word order outside of [S]VO is flexible and most of the time it is possible to change the order of peripheral elements without altering the meaning of the sentence or making it ungrammatical. In contrast, the syntax of Dutch is known to be difficult for L2 learners, and especially for Spanish learners of Dutch, exactly because Dutch has a covert SOV word order pattern, with the property known as Verb Second (V2) that only operates in main clauses (e.g., *Morgen koop ik het boek in Granada*, ‘Tomorrow I will buy the book in Granada’). At the onset of learning Dutch, adult Spanish learners tend to copy their Spanish syntax and transfer it to their Dutch sentences¹⁰. As a consequence, Spanish learners of Dutch are likely to produce ungrammatical sentences like **Morgen ik koop het boek in Granada*, ‘Tomorrow I will buy the book in Granada’. Dutch word order can get more complicated for L2 learners, for example,

⁹ Subject pronouns are commonly omitted in Spanish because verbal morphemes convey the necessary information to identify the subject of the sentence.

¹⁰ This assumption, although not attested by empirical studies, is based on the experience of teachers of Dutch as a second language at Radboud in'to Languages, the language learning centre of the Radboud University Nijmegen.

when the finite verb is a separable verb (e.g., *goedkeuren*, ‘to approve’, *afkeuren*, ‘to reject’). That is, L2 learners, and particularly Spanish learners often have problems in finding the correct word order for the particle of separable verbs in sentences like *Ik keur dit project goed*, ‘I approve this project’, *Ik heb het project goedgekeurd*, ‘I have approved the project’ or *Ik zal dit project goedkeuren*, ‘I will approve this project’.

Another fundamental difference between Spanish and Dutch is the syllable structure. Spanish tends towards an open syllable structure (CV), unlike Dutch, English and German, which tend to have a closed structure (CVC) (Booij, 1995; Hualde, 2005; see also Tropic (1987) for the sonority hierarchy within the syllable and its influence on L2 phonology). Spanish is a syllable-timed language (‘sounds if’ the duration of every syllable is constant) and Dutch a stressed-timed language (‘sounds if’ the interval between two stressed syllables is constant), terms which refer to the auditory impression produced by the language. Prosodic structure differs across languages and can also influence the perception of non-native speech by native listeners (Cutler, 2012)¹¹. To the Dutch ear, for instance, the rhythm of Spanish-accented Dutch can sound monotonous, similar to a machine gun-like sound sequence, as Spanish learners of Dutch have a tendency to transfer their Spanish rhythm (all syllables, stressed and unstressed, have the same duration) to Dutch. Along similar lines, the rhythm of Dutch-accented Spanish can sound divergent to Spanish ears, like a Morse code-like sound, most likely because the differences in duration between stressed and unstressed syllables are greater in Dutch than in Spanish (cf. Nespor, Shukla, & Mehler, 2011).

¹¹ See also Van Maastricht, Kraemer, and Swerts (2016b) who investigated how prosodic deviance by native and non-native speakers (both Spanish and Dutch) affects native speaker perceptions in terms of accentedness, intelligibility, comprehensibility and nativeness.

1.4 The phonology of Spanish¹² and Dutch

When describing the phonology of a language, it is customary to describe its segmental and suprasegmental¹³ features. In this chapter, only the segmental features (vowels and consonants) of Spanish and Dutch are discussed, as this is the focus of the present investigation.

1.4.1 Vowels

The phonological properties of the Spanish and Dutch vowel systems are very different. The most obvious difference between the Spanish and Dutch vowel systems is that Spanish has a straightforward five-vowel system (/a, e, i, o, u/; see Figure 1.1) (Hammond, 2001; Hualde, 2005; Quilis & Fernández, 1985), whereas Dutch has a complex, large-sized inventory of 15 full vowels (monophthongs: /i, y, u, ɪ, ɔ, ε, ɑ, a:/; long mid vowels: /e:, ø:, o:/; diphthongs: /ei, œy, ou/; see Figure 1.1), next to the reduced vowel /ə/ (Adank, Van Hout, & Smits, 2004b; Booij, 1995; Gussenhoven, 1999). Four features characterize the differences between the two vowel systems, as presented in Table 1.1. Spanish does not have phonemic vowel length (Hualde, 2005; McAllister, Flege, & Piske, 2002), whereas Dutch has a strict lax/tense distinction (lax vowels: /ɪ, ε, ɔ, ʏ, ɑ/; tense vowels: /i, y, u, e:, ø:, o:, a:/), which crosses the

¹² We are aware of the phonetic differences among the geographical varieties of the Spanish language in Spain and in Latin America (see Hualde (2005) for a detailed description of the Spanish language in Spain and in Latin America). In this dissertation we focus on the vocalic and consonantal phonemes of Standard Spanish.

¹³ See also Van Maastricht, Kraemer, and Swerts (2016a) for prominence patterns and prosodic transfer from Spanish to Dutch and vice versa.

short/long distinction¹⁴ (short vowels: /i, y, u, ɪ, ʏ, ə, ε, a/; long vowels: /a:, e:, ø:, o:, ei, œy, ou/) (Adank et al., 2004b). Dutch has four lax/tense vowel pairs, namely, /ɪ/-/e:/, /ʏ/-/ø:/, /ə/-/o:/, /a/-/a:/, whose lax vowels are short and whose tense vowels are long. The durational contrast in these vowel pairs correlates with a difference in the position of the tongue root, which is advanced somewhat during the pronunciation of the long, or tense, vowels (e:, ø:, o:, a:), while it is not advanced during the pronunciation of their short, or lax, counterparts (ɪ, ʏ, ə, a). In addition, the position of the body of the tongue tends to be higher for tense vowels due to “pharyngeal expansion”, i.e., tongue-root advancement (ATR), a change that does not take place in lax vowels, which are characterized by a tongue-root retraction (RTR) (Gussenhoven & Jacobs, 2011: 30; see also Botma and Van Oostendorp (2012: 141–145) for a detailed explanation on the lax/tense distinction in Dutch). Spanish does not have the feature of front rounding, as all rounded vowels in Spanish are back vowels (/o, u/) (Hualde, 2005), whereas Dutch has four front rounded vowels (/y, y, ø:, œy/). Spanish does not have diphthongs at the phoneme level, that is, single phonemes defined by their trajectory between two vowel positions; instead it has a rich inventory of 14 vowel combinations (/ie, ei, ia, ai, io, oi, iu, ui, ua, au, ue, eu, uo, ou/) (Hualde, 2005: 79). Dutch, on the other hand, has diphthongs at the phoneme level, such as /ei, œy, ou/. The Dutch long mid vowels (/e:, ø:, o:/) are not considered to be full diphthongs, but they are slightly diphthongized (cf. Adank et al., 2004b; Van der Harst, Van de Velde, & Van Hout, 2014). Finally, Spanish

¹⁴ Although the Dutch high tense vowels /i, y, u/ are phonetically short, they have longer realizations in specific phonetic contexts (e.g., before /r/, as in *duur* /dyr/, ‘expensive’) and, in particular, in words of foreign origin (Booij, 1995: 15–16.)

distinguishes three height values (high: /i, u/; mid: /e, o/; low: /a/) (Hualde, 2005), whereas Dutch is characterized by four height values (high: /i, y, u/; high mid: /ɪ, ʏ, e:, ø:, o:/; low mid: /ɔ, ɛ, ɛi, œy, ɔu /; low: /ɑ, a:/) (Booij, 1995), as shown in Figure 1.1.

Figure 1.1 makes clear that, acoustically speaking, all Spanish vowels are located at the periphery of the F1/F2 vowel space, whereas Dutch vowels also occupy the central area of the vowel spectrum (cf. Goudbeek, Cutler, & Smits, 2008).

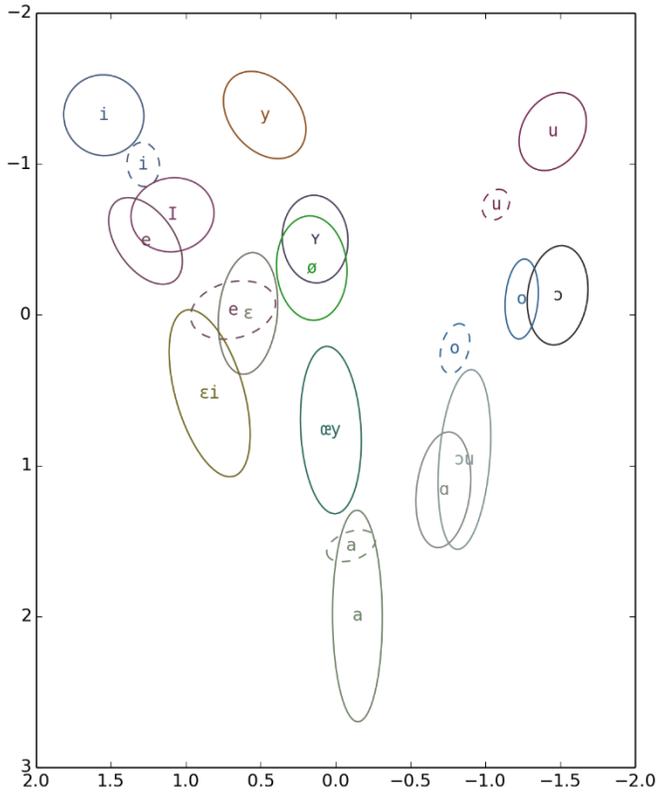


Figure 1.1 $F1$ (y -axis) and $F2$ (x -axis) (normalized) values for the five Spanish vowels (dashed lines), as realized by native Spanish speakers and for the fifteen Dutch vowels, as realized by native Dutch speakers; all vowels are measured at 50% of the vowel duration; the mean values are indicated by the vowel symbols; the values for Spanish vowels were drawn from Chládková, Escudero, and Boersma (2011); the values for Dutch vowels were drawn from Van der Harst (2011).

Table 1.1 *Distinctive features of the Spanish and Dutch vowel systems.*

	Spanish	Dutch	Comment
Length (duration)	–	+	Dutch has a complex interaction between the distinctions of lax/tense and short/long.
Front rounding	–	+	Both languages have back rounded vowels.
Diphthong	–	+	Spanish has 14 vowel combinations, but no diphthongs at the phoneme level.
Height	3 levels	4 levels	Dutch has an additional height value, as can be observed in the vowel space of Figure 1.1.

1.4.2 Consonants

Spanish and Dutch share many consonantal sounds, but not the glottal /h/¹⁵ and the labiodental /v/, which are found in Dutch, and the interdental fricative /θ/, the prepalatal affricate /tʃ/ and the rhotic trill /r/, which occur in Spanish (cf. Hualde (2005) and Booij (1995) for a detailed description of the consonantal phonemes of Spanish and Dutch respectively).

Spanish does not have as many consonant clusters as Dutch. As mentioned earlier, Spanish has a tendency for an open syllable structure (CV) (Hualde, 2005), whereas the preference in Dutch is for a closed one (CVC) (Booij, 1995). To understand the differences in syllable structure and syllabification in Spanish and Dutch, it is helpful to refer to the notion of sonority. The phonemes of a language can be arranged along a scale of sonority from more open or vowel-like to more closed or consonant-like. Every language has its own scale of sonority or sonority hierarchy, and consonant clusters which can occur in a certain language may be in violation with the sonority hierarchy of another language (see also Clements (1990), Parker (2002, 2012) and Zec (1995) for more information on the sonority hierarchy). Tropic (1987) investigated the production of German consonant clusters by Spanish

¹⁵ The grapheme <h> is silent in Spanish.

subjects¹⁶ to provide insights into the role of sonority in the variability of L2 phonology acquisition. His findings revealed that, in a given context, the ease with which certain (single) consonants and consonant clusters are acquired correlates with their degree of sonority in the subjects' native language.

The fact that Spanish has fewer consonant clusters than Dutch is due to restrictions in its sonority hierarchy (cf. Hualde (2005: 72) for an explanation of the sonority hierarchy in Spanish). Onset clusters in Spanish are always in the sequence [plosive] or /f/ + /l/ or /l/ (e.g., /pʎ/ as in *pri-me-ro*, 'first', /bʎ/ as in *bro-ma*, 'joke', /tʎ/ as in *tren*, 'train', /pʎ/ as in *pla-ne-ta*, 'planet', /bʎ/ as in *blan-co*, 'white' or /ʎ/ as in *flor*, 'flower'). Spanish codas are more restricted than Dutch codas, as they contain just one or a maximum of two consonants, of which the last is always /s/ (e.g., /ns/ as in *ins-truc-tor*, 'instructor') (Hualde, 2005). Dutch, like other Germanic languages such as English¹⁷ or German, has multiple combinations for onset and coda clusters which do not exist in Spanish (cf. Booij (1995) for a detailed description of consonant clusters in Dutch). The Dutch onset clusters /kn/ as in *knuffel*, 'hug' or /sl/ as in *sloom*, 'slow', and coda clusters /lk/ as in *melk*, 'milk' or /ts/ as in *fiets*, 'bike', which do not occur in Spanish, are known to be especially difficult for Spanish learners of Dutch. To our knowledge, there are no studies addressing the issue of the production of clusters by Spanish learners of Dutch (apart from the present investigation; see Chapter 2). Several studies addressing cluster

¹⁶ Tropf's subjects were Spanish guest workers who arrived in Germany with the first wave of migrants in the 1960s. In his doctoral dissertation, Tropf investigated the variation in the phonology of these Spanish guest workers who acquired German in an untutored L2 environment (cf. Tropf, 1983).

¹⁷ See also Ernestus, Kouwenhoven, and Van Mulken (2017) for difficulties in the comprehension of English *can* and *can't* by native Spanish listeners due to phonotactic constraints in their native language.

acquisition by Spanish learners of German (Tropf, 1987) and English (Carlisle, 1991; Yavaş & Someillan, 2005) have concluded that the combinations ‘s + stops’ (e.g., /sp/ as in *Spanje*, ‘Spain’, /st/ as in *station*, ‘station’, /sk/ as in *skelet*, ‘skeleton’) are particularly challenging for native Spanish speakers because Spanish words cannot start with a /sC/ cluster, which often leads to vowel epenthesis (sC ↦ esC).

1.5 Pronunciation problems of Spanish learners of Dutch, English, German and French

This section focuses on the problems adult Spanish learners have in acquiring an L2. Spanish learners find Dutch pronunciation problematic and some of their mispronunciations can lead to hilarious – and sometimes embarrassing – misunderstandings. For example, this researcher’s Spanish friends have found themselves in uncomfortable situations when they asked for **viesfrietjes* /visfritjəs/, ‘dirty French fries’ (instead of *visfrietjes* /vɪsfritjəs/, ‘fish French fries’) at a supermarket or when ordering a *borst* /bɔrst/¹⁸, ‘breast’ (instead of a *worst* /vɔrst/, ‘sausage’) at a food stand.

Outside of anecdotal examples, however, little is known about the pronunciation difficulties of Spanish learners of Dutch. The great majority of studies on the Spanish L1-Dutch L2 pair have focused on vowel perception. They have addressed Dutch vowel perception by naïve Spanish listeners (Escudero & Williams, 2011), by Spanish listeners with limited exposure to Dutch (Goudbeek, et al., 2008; Escudero, 2015), and by Spanish L2 learners

¹⁸ The Dutch labiodental phoneme /v/, which is represented by the grapheme <w>, is often mispronounced as /b/ because in Spanish the grapheme <w> in syllable-initial position is often realized as /b/.

(Escudero, Benders, & Lipski, 2009; Escudero & Wanrooij, 2010; Escudero & Williams, 2012; Escudero, Simon, & Mulak, 2014). Escudero and Williams (2011) studied cross-language categorization of Dutch vowels by naïve Spanish listeners and observed that five Dutch monophthongs /i, u, ɔ, ε, a:/ and two Dutch long mid vowels /e:, o:/ were assimilated primarily to a single Spanish vowel category /i, u, o, e, a/ (i.e., the five Spanish core vowels) or to a single Spanish vowel combination /ei, ou/ respectively. Other vowel tokens of Dutch /I, y, ʁ, ɑ, ø:/ were categorized in terms of two or more Spanish vowel categories, namely /i/ or /e/, /i/ or /u/, /e/ or /u/, /a/ or /o/ and /e/ or /ei/ or /eu/ respectively. In a similar vein, Goudbeek et al. (2008) investigated the acquisition of three novel Dutch phonetic categories /y, ʁ, ø:/, all three vowels being front and rounded, by Spanish listeners with limited exposure to Dutch. In their experiment, the distributional properties of the input (duration and vowel height) and the availability of supervision (supervised learning vs. unsupervised learning) were varied across several conditions presented to the participants. Their findings revealed that for the novel vowels /y, ʁ, ø:/, Spanish listeners resorted more often to F1 information (vowel height) to categorize the Dutch novel contrast /ʁ/-/ø:/ (distinguished primarily by duration in native Dutch), and that this categorization was superior when supervised learning was employed. More importantly, Goudbeek et al. (2008) concluded that Spanish learners, and L2 learners in general, find it extremely difficult to simultaneously use more than one cue to make a contrast (cf. Cutler, 2012). However, not all novel Dutch vowels are categorized by Spanish listeners on the basis of vowel height. Escudero et al. (2009) showed that duration (instead of spectral cues) was the primary perceptual cue for Spanish learners when categorizing the Dutch vowel contrast /ɑ/-/a:/ (based on both duration and vowel height in native Dutch). These findings confirm

Goudbeek et al.'s (2008) results in showing that Spanish learners have difficulty simultaneously applying more than one cue to a contrast. The same Dutch vowel contrast, namely /ɑ/-/ɑ:/, was found to be the most difficult vowel distinction to perceive for Spanish learners¹⁹. Findings by Escudero and Williams (2012) revealed that, in a categorical discrimination task and a forced-choice identification task, the Dutch vowel contrasts /ɑ/-/ɑ:/, followed by /ɪ/-/i/, was the most difficult to discriminate, most likely because the two L2 phones in each pair are non-contrastive in Spanish as both resemble the Spanish /a/ and /i/ respectively, as shown in Figure 1.4.



Figure 1.4 *Most difficult perceptual Dutch contrasts for Spanish learners, as reported by Escudero and Williams (2012).*

Other studies on the Spanish L1-Dutch L2 pair have focused on the effect of Spanish orthography on the perception of Dutch vowels. Spanish has a transparent orthography, i.e., the phoneme-grapheme correspondence is straightforward (one phoneme tends to correspond to one grapheme only),

¹⁹ See Wanrooij, Escudero, and Raijmakers (2013) who showed that distributional training helps Spanish learners to improve their perception of the difficult Dutch vowel contrast /ɑ/-/ɑ:/, and Wanrooij and Boersma (2013) who demonstrated that both continuous and discontinuous bimodal distribution of this contrast can be used for distributional learning experiments. See also Wanrooij (2015) for more information on distributional learning of vowel categories in infants and adults.

whereas Dutch has a deeper orthography, in which the phoneme-grapheme correspondence is less clear-cut (one grapheme can represent more than one phoneme, and more than one grapheme can stand for a single phoneme) (cf. Escudero & Wanrooij, 2010). Research has shown that orthography has an effect on Dutch L2 vowel learning by Spanish learners (Escudero & Wanrooij, 2010; Escudero et al., 2014; Escudero, 2015). Learning Dutch L2 vowels appears to be impeded when the Dutch spelling conventions do not match the grapheme-phoneme correspondence in Spanish (Escudero et al., 2014), especially in the case of the perceptually difficult contrasts /a/-/a:/ and /i/-/i:/ (Escudero, 2015). Along similar lines, the use of digraphs in Dutch, such as <uu>, <aa>, <ee>, <oo> to represent the Dutch vowels /y, a:, e:, o:/, might induce lengthening of these vowels in Spanish learners who are not familiar with such digraphs in their L1 orthography (Escudero & Wanrooij, 2010; Escudero et al., 2014).

While only a modest number of studies have investigated the difficulties of Spanish learners of Dutch, and most of them have focussed on L2 vowel perception, extensive research has been conducted on the difficulties Spanish learners encounter when learning English²⁰. Researchers have studied difficulties related to the perception and production of English vowels (e.g., Escudero, 2006; Escudero & Boersma, 2004; Flege, 1991; Flege & Bohn, 1989; Morrison, 2006, 2008, 2009) and consonants (e.g., Flege & Eefting, 1987), including consonant clusters (e.g., Carlisle, 1991, Yavaş & Someillan, 2005). When it comes to the acquisition of English vowels and contrasts, most scholars agree that various difficulties related to vowel height, vowel length and the lax/tense distinction arise when Spanish learners have to rearrange their 5-vowel system to fit the 15-vowel English system (cf. Flege, 1995), and

²⁰ See also Kouwenhoven (2016) for register variation, discourse management and pronunciation in Spanish English.

that the English /l/-/i/ contrast (based on vowel height) is one of the most difficult to perceive and to produce (cf. Morrison, 2006). As to consonants and consonant clusters, English stops are found to be particularly challenging for Spanish learners (Flege & Eefting, 1987), although most difficulties are found in the English consonant clusters, which are responsible for a considerable number of insertions, substitutions and deletions (Carlisle, 1991; Yavaş & Someillan, 2005).

Hardly any studies have investigated the pronunciation problems of adult Spanish learners of German, another Germanic language that shares phonological properties with Dutch. One such study is Tropsch (1987), who found that adult Spanish learners have difficulties with the production of consonant clusters. Although no studies have addressed the production of German vowels by adult Spanish learners, we expect that most vowel pronunciation errors would be related to the feature of front rounding. Ulbrich (in preparation) appears to support our expectations, as preliminary findings indicate that Spanish learners have difficulties in producing the German front rounded vowels. Front rounding also occurs in French, a Romance language which is close to Spanish in terms of lexical items and morphosyntactic structures, and far when it comes to phonological properties. No studies exist on the pronunciation problems of adult Spanish learners of French, but we expect that Spanish learners would have problems with front rounding in French, just as they seem to do in German (Ulbrich, in preparation).

What happens when a native speaker of Spanish (L1) who has prior linguistic knowledge of another Germanic language like English (L2) decides to learn Dutch (L3 or additional language (La))? It is known that prior linguistic knowledge in multilinguals can be used during their acquisition of an La (De Angelis, 2007). According to Schepens, Van der Slik, and Van Hout (2016), this prior linguistic knowledge can affect the learnability of an La, and

the closer the L2 is to the La, the higher the learnability of an La. The learnability of an La can be measured in terms of the number of new sounds which need to be learned. The L2-La phonological distance effect is weaker than the L1-La phonological distance effect (Schepens, 2015). What happens when speakers of Spanish with prior linguistic knowledge of English (L2) and/or German and/or French (L3 or La) decide to learn Dutch (La)? We expect that this prior linguistic knowledge will contribute to the learnability of Dutch, as learners in general are likely to apply features they have learnt from other languages in any subsequent languages they learn. So for example, we expect adult Spanish learners who are already fluent in German and/or French to be able to apply the feature of front rounding, transferred from German and/ or French, to produce front rounded vowels in Dutch. However, regardless of prior linguistic knowledge, Dutch is expected to always pose difficulties to some extent for Spanish learners because of the inherent phonological distance between Spanish and Dutch.

1.6 Speech perception models and Dutch L2 perception of Spanish learners

The extent to which L1-L2 interference affects the acquisition of L2 segments has been extensively studied in L2 speech research. The models on L2 speech acquisition that have been developed in this field are perception models. There are no models which are explicitly focused on production because scholars in Second Language Acquisition (SLA) tend to assume that perception necessarily precedes production when it comes to the acquisition of L2 speech sounds.

That is, as formulated by Flege (2005b) in his “Doom” hypothesis²¹, perceptual “attunement” to the language specific phonetic properties of heard speech (in the L1 and L2) occurs first, subsequently, perceptual representations are formed, and, eventually, production aligns to these perceptual representations.

Three speech perception models, the Speech Learning Model (SLM; Flege, 1995, 1999, 2003), the Perceptual Assimilation Model (PAM; Best, 1995) and its extension, the PAM-L2 (Best & Tyler, 2007), and the Second Language Linguistic Perception Model (L2LP; Escudero, 2005), have tried to explain learners’ difficulties in mastering the L2 phonological system in terms of perceived similarity between L1 and L2 segments. While Flege’s (1995) SLM focuses on learning and predicts constraints in the perception, and eventually in the production, of separate L2 segments by L2 learners, Best’s (1995) PAM classifies listeners’ difficulties in the perceptual assimilability of non-native segmental contrasts to native categories, whereas its extension, the PAM-L2 (Best & Tyler, 2007), focuses on the perception of L2 contrasts. Escudero (2005) has proposed an alternative perceptual model, the L2LP (see also Van Leussen and Escudero (2015) for a revision of the L2LP model), which aims to predict and explain acquisition processes in L2 speech perception based on creating L2 segmental contrasts. These models will be discussed briefly below.

Flege (2005a) states that “the primary aim of the SLM is to account for variation in the extent to which individuals learn – or fail to learn – to

²¹ The “Doom”(no plasticity) hypothesis (Flege, 2005b) holds that late/adult learners are unable to acquire the phonology of a second language in a native-like manner (cf. Scovel, 1988) See Flege’s slides for his lecture “Evidence in studies examining second language speech acquisition”, presented at the the ISCA Workshop on Plasticity in Speech Perception (PSP2005), London, UK, 15–17 June 2005. Retrieved from <http://www.jimflege.com/> (date last viewed 31/08/17).

accurately perceive and produce L2 segments”²². The SLM²³ predicts learners’ difficulties in terms of an L1-L2 comparative approach based on the interaction of two mechanisms, equivalence classification and the formation of new categories. The first mechanism causes L2 learners to erroneously interpret L2 phones as equivalents (i.e., *identical* or *similar*) to their own L1 categories. As a consequence, these L2 phones may differ considerably from native productions of the same speech segments. The following two situations may arise: one L2 phoneme is matched to one L1 phoneme (e.g., there is a nearly direct match between the Dutch vowel /u/ and the Spanish /u/), or two (or even more) distinct L2 segments fall into one single L1 category (e.g., the case of the Dutch vowels /l/ and /i/ which are non-contrastive in Spanish as both resemble the Spanish /i/). However, L2 segments that are sufficiently dissimilar from any L1 category (i.e., actually perceived as *new*) may evade the process of equivalence classification. In these cases, the second mechanism may come into play, and learners might be able to establish new phonetic categories for L2 segments. This could imply that in the case of Spanish L1-Dutch L2, the Dutch vowels /l/ and /i/, which are non-contrastive in the L1 and can be considered *similar* to Spanish /i/, will pose greater difficulties to Spanish learners than the *new* front rounded vowels /y/ and /ɣ/.

²² See Flege’s (2005a) slides for his keynote lecture “Origins and development of the Speech Learning Model”, presented at the 1st ASA Workshop on L2 Speech Learning, Simon Fraser University, Vancouver, BC, Canada, 14–15 April 2005. Retrieved from <http://www.jimflege.com/> (date last viewed 31/08/17).

²³ Later studies by Flege (2009, 2012) have focused on the quality and quantity of input in second language speech learning. See also Flege’s slides for his lecture “The role of input in second language (L2) speech learning”, presented at the VIth International Conference on Native and Non-native Accents of English, Łódź, Poland, 6–8 December 2012. Retrieved from <http://www.jimflege.com/> (date last viewed 31/08/17).

which do not have a Spanish counterpart and are dissimilar from any native category.

While Flege's SLM predicts learners' difficulties in terms of cross-language comparisons of phonetic categories based on the interaction of equivalence classification and the formation of new categories, Best's (1995) PAM predicts L2 listeners' difficulties in the perceptual assimilability of six non-native segmental contrasts to native categories. According to Best (1995), L2 listeners strive to maintain contrasts between L1 and L2 phonetic categories which exist in a common phonological space. The discrimination between L2 contrasts can be hard or easy, depending on the proximity between L2 segments (in a contrast) and the proximity of these L2 contrasts to L1 categories. The six assimilation patterns for non-native contrasts addressed in PAM are as follows:

- 1) *Single-category assimilation (SC type)*: Two L2 segments that learners can judge as good exemplars of a single L1 category will be difficult to discriminate.
- 2) *Two-category assimilation (TC type)*: Two L2 segments that are assimilated to two different L1 categories will be easier to differentiate.
- 3) *Category-goodness difference (CG type)*: Contrasting L2 segments are perceived as different in their relative "goodness of fit" to a single L1 category, and their discrimination will be moderate to good. Both L2 segments are assimilated to the same L1 category, but one is perceived as more deviant than the other.
- 4) *Uncategorized versus categorized (UC type)*: One L2 segment can be assimilated to an L1 category, whereas the other falls within the phonological space and outside L1 categories. Discrimination in this case is expected to be very good.

5) *Both uncategorizable (UU type)*: Both L2 segments fall within the phonological space of the L1, but are dissimilar from any L1 category (e.g., a contrast from a large-sized vowel inventory, for L2 learners with a straightforward and restricted L1 inventory). Discrimination is expected to range from poor to very good depending on the acoustic distance between the L2 segments, and on their proximity to L1 categories within the common phonological space.

6) *Nonassimilable (NA type)*: Both L2 categories fall outside the phonological space and will be perceived as nonspeech sounds. A pair of L2 categories can vary in their discriminability as nonspeech sounds and discrimination is expected to vary from good to very good.

These contrasts cover both the situation with a more extended and a more constrained sound inventory in the L2 in comparison to the L1.

An alternative model to Flege's (1995) SLM and Best's (1995) PAM is the L2LP model (Escudero, 2005; Van Leussen & Escudero, 2015). The L2LP model postulates the Full Copying hypothesis²⁴, which states that L2 learners will initially perceive all L2 segments as exemplars of their native categories. In the initial stage of the learning process, L2 learners will copy their L1 perception and will make use of L1-learning mechanisms, aimed at developing an optimal L2 perception. The L2LP model proposes precise learning tasks and developmental paths for L2 learners, depending on the learning scenario they are faced with. The L2LP model reduces the six types of non-native contrasts addressed in Best's (1995) PAM to three prototypes. These three prototypes or learning scenarios in the L2LP are referred to as *New scenario*, *Similar scenario* and *Subset scenario*, as shown in Figure 1.5.

²⁴ Escudero's Full Copy hypothesis resembles the mechanism of equivalence classification proposed by Flege (SLM; 1995, 1999, 2003).

In the *New scenario*, L2 learners are confronted with L2 phonological phonemes that do not exist in their L1. The *New scenario* in the L2LP is what PAM (Best, 1995) calls *Single-category assimilation*. In the *New scenario* the number of target L2 categories is larger than that of L1 categories. An example of this scenario is that of the Dutch vowels /ɪ/ and /i/, which are likely to be associated by Spanish learners with the Spanish vowel /i/. In the *Similar scenario* (PAM: *Two-category assimilation*), the phones of an L2 contrast are acoustically closest to the productions of two separate L1 sounds. In this case, L2 learners are likely to simply replicate their existing L1 categories and to adjust their boundaries to fit the L2 contrast. For instance, for Spanish learners, the phones in the Dutch contrast /ɪ/-/ɛ/ are acoustically similar to the phones in the L1 contrast /i/-/e/. Therefore, Spanish learners will simply replicate their existing L1 phones /i/ and /e/, or will adjust the boundaries of these two L1 phones to fit the Dutch vowels /ɪ/ and /ɛ/ respectively. The L2LP predicts that shifting the boundaries of existing L1 categories will be less problematic than creating new categories altogether. In the *Subset scenario*²⁵ (PAM: *Uncategorized versus categorized*), a single L2 sound is perceived by L2 learners as being related to more than one L1 category. This scenario occurs when learners have more detailed phonetic contrasts in the L1 than in the L2, which applies, for instance, to vowel contrasts when native Dutch speakers learn Spanish. The *Subset Scenario* is not applicable to Spanish learners of Dutch, that is, to the Spanish L1-Dutch L2 pair which is the focus of this dissertation. As advanced by Escudero (2005: 123), the *Subset scenario* has not been explicitly considered in previous L2 perception models. These have commonly considered only two scenarios, i.e., similar and new (cf. Van

²⁵ According to Escudero's (2005: 123) L2LP model, L2 learners face a subset scenario because the L2 category constitutes a subset of two L1 categories.

Leussen & Escudero, 2015). According to the L2LP, the *Subset scenario* is likely to be less difficult for L2 learners than the *New scenario*, as no new contrasts have to be created. For example, the Spanish vowel /e/ is perceived by Dutch learners of Spanish as two existing L1 categories, namely, as the Dutch vowels /ɪ/ and /ɛ/. The L2LP states that learners, depending on the learning scenario they are confronted with, will perform precise learning tasks through auditory-driven category formation and lexicon-driven category boundary shifting (Escudero, 2005: 122). Eventually, exposure to the L2 will help L2 learners to gradually abandon their L1-routines and full copying state, and to enter a state which approximates L2 perception and recognition in a more native-like manner.

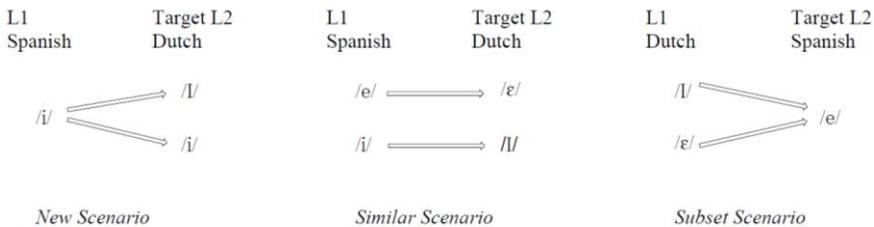


Figure 1.5 *New, Similar and Subset scenarios, as proposed in the L2LP model (Escudero, 2005; Van Leussen & Escudero, 2015). The Subset Scenario is not applicable to the Spanish L1-Dutch L2 pair.*

As we have seen, there is a clear resemblance between Flege's SLM, Best's PAM and Escudero's L2LP model. A shared tenet of these speech perception models is that L1-specific sound patterns affect L2 speech perception. This form of L1 entrenchment, which is more evident in adult than in early L2 learners, draws the learner's attention to contrastive phonetic elements or features that are relevant in the L1, possibly leading to an unfortunate situation

where they do not notice phonetic distinctions that are essential in the L2. When listening to a second language, adult L2 learners resort to their L1 listening structure to map the incoming input of L2 phonemes onto their existing L1 phoneme repertoire. The influence of the L1-listening structure can lead to phonemic misperceptions which can affect word recognition, as both lexical candidates and lexical competitors can be activated (Cutler, 2012). Well-known examples of these phonemic misperceptions by Spanish learners of Dutch can be found in the perception of minimal pair contrasts such as *man* /man/, ‘man’, *maan* /ma:n/, ‘moon’; *vis* /vɪs/, ‘fish’, *vies* /vis/, ‘dirty’; *duur* /dyr/, ‘expensive’; *deur* /dø:r/, ‘door’ or *borst* /bɔrst/, ‘breast’, *worst* /vɔrst/, ‘sausage’. Spanish learners, who listen to Dutch sounds with Spanish ears, require fine-grained perceptual properties to recognize these contrasts (i.e., /ɑ/-/a:/, /l/-/i/, /y/-/ø:/ and /b/-/v/), which are different from the phonemic contrasts they are used to in their native language.

In our analyses we focused on Flege’s SLM to predict and explain learners’ difficulties in mastering the L2 phonological system. We chose the SLM as it is focused on learning (cf. Cutler, 2012: 307–308), and its strength is explaining why some L2 segments are more difficult for adult L2 learners to learn than others, both in perception and eventually production, although, like the other models, SLM does not predict the full range of variability in the production of L2 segments. Wherever helpful we used the L2LP model to predict difficulties and to interpret findings, as this model emphasizes the role of determining contrastive values or features between segments. An important question is which Dutch vowel contrasts are interpreted as new and which contrasts tend to be interpreted as equivalent, and how and to what extent Spanish learners eventually solve incorrect equivalence classifications in their production.

1.7 Perception of foreign-accented speech by native listeners

Cutler (2012) concludes that native listening is an extremely intricate process, by which information cascades are constantly being weighed, allowing higher language-specific probabilities to influence phonemic and lexical decisions. Native listeners are exceptional in their ability to rapidly adapt to (new) speech: new words, pronunciation variability across talkers, dialectal variation and unfamiliar accents. Native listeners' perceptual adaptation rests on the plasticity of adult speech perception, by which native phonemic boundaries can adapt in order to facilitate communication. Thus, a basic and universal principle of native listening is the adaptive nature of speech processing, both to the great variability in native speech and to the structure of the native language. These speech processes are governed by language-specific aspects which appear to be different across languages (Cutler, 2012).

Listening, which feels like the easiest thing to do, can become harder when listening to foreign-accented speech. When listening to foreign-accented speech, native listeners attend to phonetic details resulting from transfer from the learners' native language, and need to navigate specific types of deviations in the speech signal²⁶. Recognizing words with segmental deviations implies that listeners have to cope with sounds that are reduced versions of target phones, as well as with sounds that can be mapped onto a distinct phoneme category, which may cause confusion with other words (cf. Bent, Baesse-Berk, Borrie, & McKee, 2016; Cutler, 2012).

For native Dutch listeners who listen to Spanish-accented Dutch, poor L1-L2 sound mapping can lead to phonemic misperceptions, which can affect word recognition, as both lexical candidates and lexical competitors can be

²⁶ See Witteman, Weber, and McQueen (2014) who showed that native Dutch listeners can tolerate inconsistency in German-accented Dutch and can rapidly adapt to the speaker.

activated (Cutler, 2012). Imagine an adult Spanish learner producing the Dutch words mentioned earlier (i.e., *man* /man/, ‘man’, *maan* /ma:n/, ‘moon’; *vis* /vɪs/, ‘fish’, *vies* /vis/, ‘dirty’; *duur* /dyr/, ‘expensive’; *deur* /dø:r/, ‘door’ or *borst* /bɔrst/, ‘breast’, *worst* /vɔrst/, ‘sausage’), namely those words containing the phonemic contrasts /ɑ/-/a:/, /l/-/li/, /y/-/ø:/ and /b/-/v/. How are these phonemic contrasts produced by Spanish learners of Dutch perceived by native Dutch listeners? Native Dutch listeners, like Spanish learners, filter their perception of Spanish-accented Dutch through their L1-sieve, and can easily detect mispronunciations which differ from their experience with target phoneme categories (Cutler, 2012), especially mispronunciations that native listeners are not likely to produce (Magen, 1998). The variability in Spanish learners’ productions requires effort from the native Dutch listeners as they have to adapt to different pronunciations, as well as different pronunciation errors, which may also differ widely across individual learners. Some pronunciation errors or modifications in the realization of phonemic vowel categories may result in native listeners having to shift their category boundaries to accommodate an ambiguous vowel realization that differs from their usual expectations about phonemic categories (Cutler, 2012). These perceptual adaptation processes, by which native listeners adapt to foreign-accented speech, show that adaptation in the L1 is not fixed, as native listeners can adjust their boundary shift between categories to a context with great variability in L2 speech across individual learners whenever language processing can benefit as a result (cf. Bradlow & Bent, 2008; Cutler, 2012).

1.8 Research objective and design

The main aim of this investigation was to identify the most frequent segmental errors of Spanish learners' productions and their possible sources, and how these learner productions are perceived by native Dutch listeners, in terms of intelligibility. To this end, four main research questions were formulated:

RQ₁: *What are the most frequent segmental pronunciation problems of adult Spanish learners of Dutch, and what are the sources for these pronunciation problems?*

RQ₂: *Do the Dutch vowels produced by adult Spanish learners acoustically match those of native Dutch speakers?*

RQ₃: *Are the Dutch vowels as produced by Spanish learners of Dutch intelligible for non-expert native Dutch listeners?*

RQ₄: *Do the acoustic properties of the Dutch vowels spoken by adult Spanish learners of Dutch match the perceptual assessments by natives of these learner vowel productions?*

Chapter 2 focuses on answering RQ₁, namely *What are the most frequent pronunciation problems of adult Spanish learners of Dutch, and what are the sources for these pronunciation problems?* To this aim, a corpus consisting of recordings of spontaneous speech produced by 23 Spanish learners of Dutch was compiled (Corpus I: corpus Spanish L1-Dutch L2 spontaneous speech). All recordings were speaking exercises of official oral exams (for CEFR²⁷

²⁷ CEFR stands for Common European Framework of Reference for Languages: Learning, Teaching, Assessment (Council of Europe).

proficiency levels A1, A2, B1) and a state exam (for B2) in Dutch. The speech data was orthographically transcribed and annotated. The annotations were used to generate confusion matrices comparing the automatically generated transcriptions with the manually corrected phonemic transcriptions.

Chapter 3 aims to answer RQ₂, i.e., *Do the Dutch vowels produced by adult Spanish learners acoustically match those of native Dutch speakers?* It reports on an investigation in which we acoustically analyzed the Dutch vowels produced by adult Spanish learners and compared the learner realizations to native productions of the same Dutch target vowels. The corpus containing the native Dutch speech material was already available (see Van der Harst (2001: 56) for a detailed description of the corpus of native Dutch speech, which contains read and spontaneous speech, and varies from separate words to sentences, paragraphs and free speech). However, we needed a new Spanish L1-Dutch L2 corpus, in which the same separate words, sentences and paragraphs used in the native Dutch speech data would be recorded by Spanish learners. Therefore, we collected a more focused corpus of Spanish L1-Dutch L2 than the corpus of spontaneous speech (i.e., Corpus I) described in Chapter 2. This new corpus contained read speech (separate words, sentences and paragraphs) and sufficient productions of Dutch speech sounds that are problematic for Spanish learners (Corpus II: corpus Spanish L1-Dutch L2 read speech). For our study, we used a subset of this corpus in which Spanish learners of Dutch read monosyllabic words from a computer screen.

Chapters 4 and 5 investigate the perception of Spanish-accented Dutch by native Dutch listeners in terms of intelligibility, and aim to answer RQ₃, namely, *Are the Dutch vowels as produced by Spanish learners of Dutch intelligible for non-expert native Dutch listeners?* Chapter 4 describes a crowdsourcing study in which the same Dutch monosyllabic words produced by Spanish learners, and acoustically analyzed in Chapter 3, were employed as

speech stimuli and were orthographically transcribed by native Dutch listeners to assess the intelligibility of the learner productions. A match between the vowel transcribed by the native listener and the canonical (target) form of the same vowel would indicate that the intended vowel, as realized by the Spanish learners, was intelligible for the native Dutch listeners. The aim of the crowdsourcing study was to investigate how the *auris populi*, the crowd's ear, would deal with possibly deviant L2 vowel realizations. With a view to recruiting more participants, a game element was introduced. After completing the transcription task, participants received a score that could be shared on Facebook. Nearly 200 native Dutch listeners participated in the transcription task. A data set from these native Dutch listeners, including native transcriptions of Dutch monosyllabic words spoken by adult Spanish learners of Dutch, was compiled (Corpus III: corpus native transcriptions of Spanish L1-Dutch L2 read Dutch monosyllabic words, using crowdsourcing sampling).

The crowdsourcing study had potential drawbacks, such as the lack of direct control over the selection of transcribers and the inclusion of the game element. For this reason, a more controlled study was conducted to determine the consistency of the outcomes in sampling diverse and large groups of non-expert native listeners transcribing non-native speech. In this study, reported in Chapter 5, we used snowball sampling, which consists of recruiting a large number of subjects from the social networks of a small starting set of individuals. In this follow-up study, we included native Dutch speech samples (cf. Van der Harst, 2011) which could be used as anchor points by the transcribers. The study using snowball sampling yielded transcription data from 132 non-expert native Dutch listeners who transcribed the same Dutch monosyllabic words spoken by adult Spanish learners and investigated in

Chapter 3 and Chapter 4 (Corpus IV: corpus native transcriptions of Spanish L1-Dutch L2 read Dutch monosyllabic words, using snowball sampling).

Chapter 6 aims to answer RQ₄, namely, *Do the acoustic properties of the Dutch vowels spoken by adult Spanish learners of Dutch match the perceptual assessments by natives of these learner vowel productions?* Statistical vowel classifications obtained from acoustic properties (acoustic data from Corpus II) were compared with classifications obtained from native Dutch listeners (listener data from Corpus IV). We considered the outcomes of the production and native perception data in the context of the learners' CEFR proficiency levels in Dutch, their multilingual background, length of residence and use of Dutch.

1.9 Outline of the dissertation

This section provides an overview of the remaining chapters in this dissertation.

Chapter 2 seeks to find out what the most frequent pronunciation errors of Spanish learners of Dutch are and their possible sources. To this end, recordings of extemporaneous speech produced by 23 Spanish learners of Dutch were analyzed to get insight into their pronunciation difficulties. Our findings show that among Spanish learners of Dutch, vowel errors are more frequent, persistent and variable than consonant mispronunciations. Spanish learners appear to have problems with contrasts in vowel length and vowel height, and in producing front rounded vowels. Consonant mispronunciations are found primarily in clusters, which are responsible for a considerable number of insertions, substitutions and deletions. Mispronunciations due to orthographic interference are observed for both vowels and consonants.

Chapter 3 deals with the production accuracy of Dutch vowels by Spanish learners, as it was found that most mispronunciations were related to vowels.

We collected a new corpus based on read speech which contained systematic productions of Dutch speech sounds that are problematic for Spanish learners of Dutch. Elicited, read speech from learners with varying proficiency levels was segmented and acoustically analyzed to determine the vowel space distributions and durations of their vowel productions in comparison to those of native Dutch speakers. Our findings show that the learner realizations did not match those of native speakers for duration and spectral values. This is probably due to L1 entrenchment effects. Producing Dutch monophthongs is hard for Spanish learners, particularly when vowel contrasts reflect subtle spectral differences. Therefore, they often erroneously resort to duration to realize such contrasts. In contrast, the Spanish learners were found to be successful in making the short/long distinction and in producing Dutch long mid vowels and diphthongs. Remarkably, they were also able to create a new vowel category (front round).

Chapter 4 reports on a study in which Dutch vowels produced by adult Spanish learners were orthographically transcribed by non-expert native Dutch listeners through crowdsourcing. The aim of the crowdsourcing study was to investigate how the *auris populi*, the crowd's ear, would deal with possibly deviant L2 vowel realizations. The results indicate that Dutch vowels pronounced by Spanish learners were transcribed differently from their canonical (target) forms by native listeners. The listeners' transcriptions confirm findings of previous research based on expert annotations of Spanish learners' vowel realizations conducted at our lab, namely, that the five Spanish vowels seem to function as "attractors" for the larger set of the Dutch vowels. In general, the results are also in line with the outcomes of acoustic measurements of the same speech material, but there are some interesting discrepancies. An interesting discrepancy between the listeners' transcriptions and the acoustic data was found with respect to duration. The transcriptions

do not indicate longer durations of the vowels in question, while objective measurements showed that the learners' vowels were longer than the corresponding ones produced by native speakers. This may suggest that native listeners "somehow" normalize duration in learner speech with little consequences for word recognition and intelligibility. We conclude Chapter 4 by formulating some evaluative remarks on the *auris populi* methodology, which is considered to be a valuable tool to collect large amounts of L2 speech transcriptions by a diverse group of non-expert native Dutch listeners.

Chapter 5 investigates how Dutch vowels produced by Spanish learners are perceived by a diverse and large group of non-expert native Dutch listeners, again by means of the crowd's ear methodology. Results showed that Dutch vowels pronounced by Spanish learners were often transcribed differently from their canonical forms. The outcomes consolidate earlier findings on the intelligibility of Spanish accented Dutch and on the speech production of Spanish learners. The vowel confusion patterns observed are consistent with the earlier study using crowdsourced sampling, supporting the usefulness of the crowd's ear approach for future L2 speech research. In addition, tentative results were found pointing to the occurrence of perceptual adaptation, by which native listeners retune their phoneme boundaries when exposed to a mix of non-native and native speech.

Chapter 6 analyzes the acoustic properties of Dutch vowels produced by adult Spanish learners and investigates how these vowels are perceived by a varied and extensive group of non-expert native Dutch listeners. Statistical vowel classifications obtained from the acoustical properties of the learner vowel realizations were compared to vowel classifications provided by native Dutch listeners. Both types of classifications were affected by the set of vowels included as stimuli, an effect caused by the large variability in Spanish learners' vowel realizations. While there were outspoken matches between the

two types of classifications, shifts were noted within and between production and perception, depending on the vowel and vowel features. We considered the variability between Spanish learners further by investigating individual patterns in the production and perception data, and linking these to the learners' proficiency level and multilingual background. We conclude that integrating production and perception data provides valuable insights into the role of different features in adult L2 learning, and how their properties actively interact in the way L2 speech is perceived. A second conclusion is that adaptive mechanisms, signalled by boundary shifts and useful in coping with variability of non-native vowel stimuli, play a role in both statistical vowel classifications (production) and human vowel recognition (perception).

Finally, Chapter 7 starts by summarizing the findings presented in the research chapters to answer the research questions formulated in 1.8 of this chapter. The most relevant issues to emerge from the research findings, namely, individual variability, adaptive mechanisms (both in statistical classification and human recognition), and L2 language proficiency and pronunciation are examined next. These issues and the research findings are then considered in the context of the speech perception models introduced in the present chapter. The limitations of this investigation and future prospects are discussed. The closing section addresses the societal relevance of my research.

2 Phonology acquisition in Spanish learners of Dutch: Error patterns in pronunciation

This chapter has been reformatted and slightly modified from:

Burgos, P., Cucchiaroni, C., Van Hout, R., & Strik, H. (2014a). Phonology acquisition in Spanish learners of Dutch: Error patterns in pronunciation. *Language Sciences*, 41, 129–142.

2.1 Introduction

Research on second language (L2) acquisition has shown that adult learners have difficulties in mastering L2 sound patterns with the ability of a native speaker (Birdsong & Molis, 2001; Long, 1990). An important limiting factor in acquiring the pronunciation of an L2 is the interference from the native language (L1). Several theories have been advanced to explain L1-L2 interference in speech processing and to predict its consequences (Best, 1995; Escudero, 2005; Flege, 1995; Strange, 2011). It is acknowledged that interference from L1 may cause segmental errors that can hinder communication, for instance by slowing down word recognition speed (Derwing & Munro, 1997; Rogers & Dalby, 1996). In addition, a foreign accent can be disadvantageous for successful interaction and social acceptance (Brennan & Brennan, 1981; Lippi-Green, 1997). Finally, many L2 learners desire to sound as native as possible eliminating traces of non-nativeness from their speech, for example because this is required for their profession.

For these reasons there is growing demand for personalized pronunciation trainings, preferably with a teacher, but this is not always feasible. This has contributed to the development of computerized pronunciation training programs, e.g., Computer Assisted Pronunciation Training (CAPT)

applications that make use of Automatic Speech Recognition (ASR), to provide sufficient practice and personalized, instantaneous feedback. An essential step in developing such programs is to gain insight into the pronunciation errors made by L2 learners. The ultimate goal of our research is to develop a dedicated CAPT program for Spanish speaking learners of Dutch L2. In order to do so, we need more information on the pronunciation errors they make. In this article we report on a study aimed at providing such information.

Previous studies investigated how Spanish L1 can affect perception of Dutch L2 (Escudero & Boersma, 2004; Escudero, Benders, & Lipski, 2009), but did not address Dutch L2 speech production by Spanish speakers. As a matter of fact, little is known about the specific pronunciation errors Spanish speaking learners make in Dutch L2. The present study aims at filling this gap by investigating Dutch L2 speech production of Spanish learners with a view to drawing up an inventory of the most frequent pronunciation errors, which can then be used as a guideline in developing dedicated ASR-based CAPT programs for this target group.

In the remainder of this article we first present general background information on research on phonological and orthographic differences that might lead to pronunciation problems in (Dutch) L2. We then compare the phonological systems of the two languages involved in our investigation, Spanish and Dutch, to identify possible sources of pronunciation difficulties. Subsequently, we describe the method, corpus design and analysis procedure of our study. The results are then presented and discussed. Finally, we draw conclusions and present future perspectives.

2.2 Research background

2.2.1 L2 speech perception models

A considerable body of research on L2 speech processing has focused on exploring how phonological differences between the L1 and the L2 can lead to difficulties in acquiring L2 speech sounds. Several models have been advanced such as the Speech Learning Model (SLM; Flege, 1995), the Perceptual Assimilation Model (PAM; Best, 1995), the Second Language Linguistic Perception Model (L2LP; Escudero, 2005) and the Automatic Selective Perception Model (ASP, Strange, 2011). The general idea is that experience with L1 and the consequent emergence of specific L1 structures leads to a form of L1 entrenchment that causes difficulties in learning to perceive and produce L2 speech sounds. In particular, it appears that L2 sound contrasts which are mapped onto an L1 single category are the most difficult to discriminate and to learn (Best, 1995; Escudero, 2005). Some researchers found that even the native dialect can influence the perception of non-native sounds (Chládková & Podlipský, 2011; Escudero, 2005; Escudero & Williams, 2012; Mayr & Escudero, 2010).

2.2.2 Orthographic interference

In addition to phonetic and phonological differences between the L1 and the L2, orthography also appears to play a role in L2 phonology acquisition. In particular L2 learners with a transparent L1 orthography, that is one in which the grapheme-phoneme correspondence is straightforward, have difficulties in acquiring the phonology of more opaque languages (Erdener & Burnham, 2005; Geva & Wang, 2001). According to Young-Scholten (2002) and Bassetti (2006) the influence of L1 orthography on L2 learning is particularly noticeable in literate adult learners who will have received written input from

the start of their exposure to the L2. Other studies have shown an effect of L1 orthography on non-native speech perception (Escudero & Wanrooij, 2010; Ortega-Llebaria, Faulkner, & Hazan, 2001).

2.2.3 *The Phonology of Spanish and Dutch*

2.2.3.1 *Vowels*. The most obvious difference between the Spanish and Dutch vowel systems is that Spanish has five vowels (/i, u, o, e, a/) (Hammond, 2001; Hualde, 2005; Quilis & Fernández, 1985) and Dutch 15 unreduced vowels (lax vowels: /ɪ, ɛ, ɔ, ʏ, ʌ/; tense vowels: /i, y, u, e:, ø:, o:, a:/; three diphthongs: /ɛi, œy, ɔu/) and the reduced vowel schwa /ə/ (Adank, Van Hout, & Smits, 2004b; Booij, 1995; Gussenhoven, 1999;). Moreover, Dutch has a lax/tense distinction and front rounded vowels: /ʏ, y, ø:, œy/, which do not exist in Spanish.

Acoustically speaking all Spanish vowels are located at the periphery of the F1/F2 vowel space, whereas Dutch vowels also occupy the central area of the vowel spectrum (Goudbeek, Cutler, & Smits, 2008). Cervera, Miralles and González-Álvarez (2001) obtained average formant values for the five Spanish vowels in a study on 10 male subjects who spoke standard Castilian Spanish. Their results are shown in Table 2.1.

Table 2.1 *Mean vowel durations (in ms) and F1 and F2 frequencies (Hz) of Spanish vowels (Cervera et al., 2001).*

	/i/	/u/	/o/	/e/	/a/
Duration	81	75	83	78	89
F1	331	376	533	502	718
F2	2241	773	1156	1872	1479

Adank et al. (2004b) present an acoustical description of all 15 vowels produced by twenty male talkers of Standard Dutch in the Netherlands. The average formant values of this study are given in Table 2.2.

Table 2.2 *Mean vowel durations (in ms) and F1 and F2 frequencies (Hz) of Dutch vowels (Adank et al., 2004b).*

	/i/	/y/	/u/	/ɪ/	/ɥ/	/ɔ/	/e/	/a/	/a:/	/e:/	/ø:/	/o:/	/ei/	/œy/	/ɔu/
Duration	94	93	98	82	88	90	95	96	203	181	184	184	192	189	195
F1	278	259	259	361	366	402	475	578	670	400	375	412	543	574	580
F2	2162	1734	805	1919	1595	821	1739	1172	1425	1995	1563	929	1804	1515	1117

Dutch has a clear duration or length distinction, the most open vowel, the mid-tense vowels and the diphthongs being long: /a:, e:, ø:, o:, ei, œy, ɔu/. This length difference correlates with systematic spectral differences in the F1 and F2 values between the lax/tense vowel pairs: /a/-/a:/, /ɪ/-/e:/, /ɔ/-/o:/, /ɥ/-/ø:/, as can be observed in Figure 2.1.

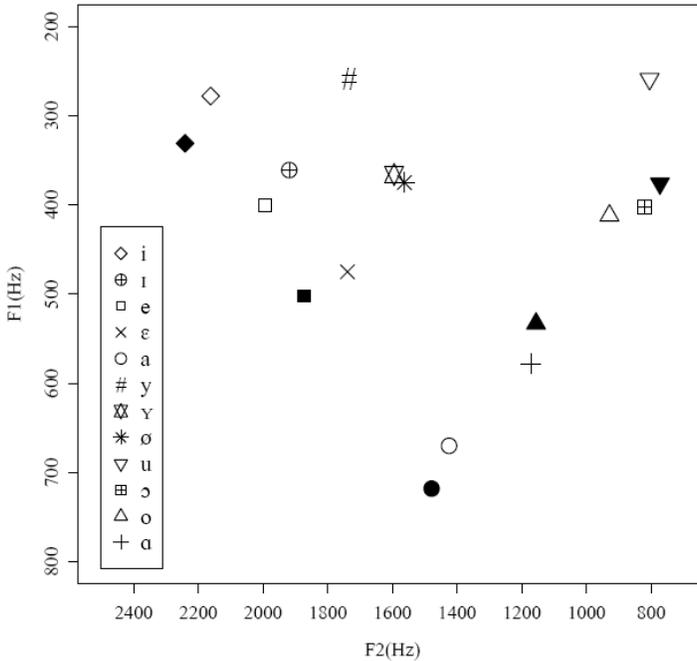


Figure 2.1 The vowel spaces of the Spanish and Dutch vowels (without the diphthongs), according to their F1 and F2 values; the Spanish vowels are represented by the filled symbols (black), the Dutch ones by the open symbols.

Figure 2.1 also shows that the Spanish vowels have lower F1 values than their Dutch counterparts (see the locations of the Dutch and Spanish /i/, /u/ and /a/). The most self-evident explanation seems to be the difference in average height between men in Spain and the Netherlands. According to The Spanish *INE* (*Instituto Nacional de Estadística*, Spanish National Institute of Statistics) the average height of Spanish adult male between 25-34 years was 175 cm in 2001, while the Dutch *CBS* (*Centraal Bureau voor de Statistiek*, Dutch National Institute of Statistics) indicated that 180 cm was the average height of adult Dutch males within the same age range in 2001. This may

imply a smaller vocal tract on average in the Spanish speaker sample, resulting in higher F1 values. Transformations are available to match differences in the position of vowel triangles (cf. Adank, Smits, & Van Hout, 2004a; Kendal & Thomas, 2010), such as the Lobanov transformation (Lobanov, 1971). We used the Lobanov transformation to match the Dutch and Spanish vowels for the average values reported in Tables 2.1 and 2.2. We did not take the Dutch diphthongs into account. The results are given in Figure 2.2.

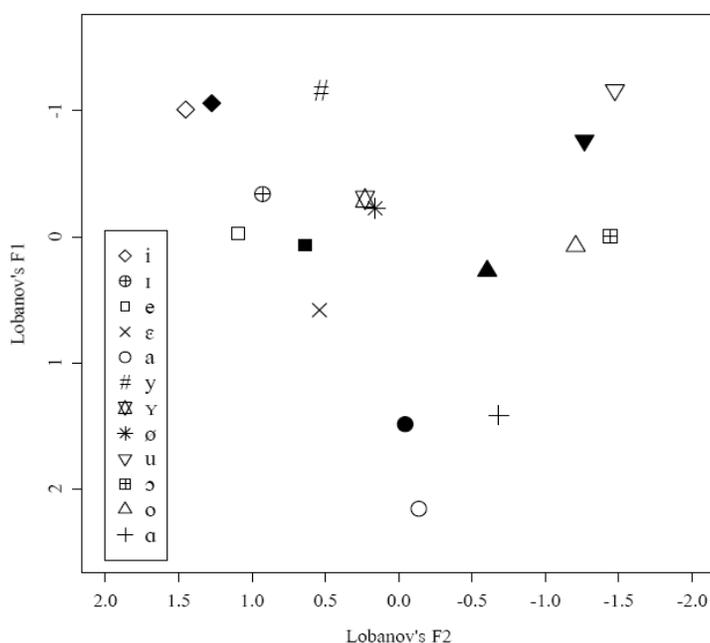


Figure 2.2 *The normalized vowel spaces of the Spanish and Dutch vowels (without the diphthongs), according to their F1 and F2 values (Lobanov transformation; Lobanov, 1971); the Spanish vowels are represented by the filled symbols (black), the Dutch ones by the open symbols.*

Figures 2.1 and 2.2 show that a mapping problem is likely to occur since Spanish learners of Dutch could map two different Dutch phonemes onto a

single Spanish category leading to possible difficulties in the discrimination and acquisition of Dutch L2 sounds as shown by Escudero and Williams (2011).

2.2.3.2 *Consonants*. Spanish and Dutch share many consonantal sounds except the glottal /h/ and the labiodental /v/, which are Dutch phonemes, and the interdental fricative /θ/, the prepalatal affricate /tʃ/ and the rhotic trill /r/ which occur in Spanish.

The pronunciation of syllable final consonants in clusters is also known to be problematic for Spanish learners, since Spanish has a tendency for an open syllable structure (CV), unlike Dutch, English or German whose preference is a closed one (CVC) (Hualde, 2005; Tropic, 1987). There are no studies addressing the issue of the production of clusters by Spanish learners of Dutch. However, several studies have investigated cluster acquisition by Spanish learners in German (Tropic, 1987) or English L2 (Carlisle, 1991; Yavaş & Someillan, 2005). Since Spanish words cannot start with a /sC/ cluster, the combinations 's + stops', i.e. /sp/ as in *Spanje*, 'Spain', /st/ as in *station*, 'station', /sk/ as in *skelet*, 'skeleton' appear to be particularly challenging for Spanish natives and vowel epenthesis (sC → esC) is often the result.

2.2.4 *Production and perception difficulties of Spanish learners of Dutch*

Little is known about the specific pronunciation errors of Spanish learners of Dutch. A restricted number of studies on the pronunciation of Dutch by students with different L1s has shed some light on this issue as Spanish was one of the native languages of the participants (cf. Neri, Cucchiari, & Strik, 2006). The most common errors concerned the vowels /ə/, which was often

deleted or substituted by the Spanish /e/; /a/ was usually substituted by the Dutch /a:/ or the Spanish /a/, the front rounded vowels /ɤ/ and /y/ were frequently mispronounced as /u/, and /ø:/ was substituted either by /y/, /o:/ or /u/. The diphthongs /ei/ and /œy/ were frequently substituted by /ei/ and /ɒu/, respectively. The Dutch consonants /t/ and /h/ were often deleted or substituted by /d/ and /x/ respectively. Mispronunciations of /x/ as either /g/, /h/ or /k/ were also noticed.

Several studies on Dutch vowel speech perception also included native Spanish speakers and can be informative in this respect. The topics addressed varied from vowel categorization (Goudbeek et al., 2008), to L2 perceptual cue weighting (Escudero, Benders, & Lipski, 2009) or to the effect of L1 orthography on non-native vowel perception (Escudero & Wanrooij, 2010). According to Goudbeek et al. (2008), Spanish listeners found it easier to categorize novel Dutch vowels by formant frequencies than by duration. Escudero et al. (2009), on the other hand, found that native Spanish listeners of Dutch L2 weight vowel duration heavier than formant frequencies in perceiving Dutch vowels. Escudero and Williams (2011) studied cross-language categorization of Dutch vowels by forty naïve Peruvian Spanish listeners and observed that the Dutch vowels /i, ε, a:, ɔ, u, e:, o:/ were assimilated primarily to Spanish /i, e, a, o, u, ei, ou/ respectively. Other vowel tokens of Dutch /I, y, ɤ, ɑ, ø:/ were categorized in terms of two or more Spanish vowel categories, namely /i-e, i-u, e-u, a-o, e-ei-eu/ respectively. Escudero and Williams (2012) studied the influence of the native dialect (Peruvian Spanish (PS) vs. Iberian Spanish (IS)) on Dutch L2 vowel perception and found that both PS and IS learners had the most difficulty with the Dutch /ɑ/-/a:/ and /I/-/i/ contrasts, but IS performed slightly better, namely with a percentage of 3%

above PS, which the authors ascribe to the specific characteristic of their native dialect.

2.2.5 Expected difficulties

Based on the results presented in the previous sections, we try to summarize which difficulties Spanish learners may experience when acquiring Dutch L2.

2.2.5.1 Vowels. Based on phonological and orthographic differences between Spanish and Dutch and on previous research on Spanish learners' Dutch L2 speech production (Neri et al., 2006) and perception (Escudero et al., 2009; Escudero & Waanrooij, 2010; Escudero & Williams, 2011, 2012; Goudbeek et al., 2008), we predict that Spanish learners of Dutch have problems in reconciling 15 unreduced vowels and the reduced vowel schwa with their own five vowel system. We expect problems in the production of the lax/tense distinction and difficulties in producing the Dutch front rounded vowels /ɤ, y, ø:, œy/.

2.2.5.2 Consonants. Since the phonemes /h/ and /v/ do not exist in Spanish, we can predict that Spanish learners of Dutch might have problems in pronouncing them. In addition, the fact that the grapheme <h> is silent in Spanish could lead to /h/ deletion in Dutch. The labiodental phoneme /v/, which in Dutch is represented by the grapheme <w>, might be mispronounced as /b/ because in Spanish the grapheme <w> in syllable-initial position is often realized as /b/. Finally, problems with clusters are also predicted which might lead to pronunciation errors such as assimilation, epenthesis or elision.

2.3 Method

2.3.1 Participants

The participants involved in this research were 5 adult male and 18 adult female native speakers of Spanish, namely 11 natives of Iberian Spanish and 12 natives of Latin American Spanish¹, who were living in the Netherlands at the time of the recordings and all reported having learnt Dutch for at least a few months. The recordings participants delivered were samples of speech which were taken at one given point in time during their Dutch learning process, which means that the data are not longitudinal. All informants were studying Dutch at Radboud in'to Languages, the language learning centre of the Radboud University Nijmegen, and were placed in a course based on their proficiency level (A1, n=4; A2, n=8; B1, n=6; B2, n=5) according to the Common European Framework of Reference for Languages (CEFR; Council of Europe).

2.3.2 Speech Material

A spoken corpus with recordings of Spanish learners of Dutch L2 was compiled. All recordings, examples of extemporaneous speech, were speaking exercises of official oral exams (for A1, A2 and B1) and a state exam (for B2) in Dutch as a second language.

It is important to underline that the choice for this type of material adds to the realistic character of our research. Our speech material is relatively natural

¹ We are aware of the acoustic differences between these two varieties of Spanish and their possible influences on Dutch L2 perception (Escudero & Williams, 2012). However, since the differences in perception are marginal, they do not warrant treating these two varieties separately in a study on Dutch L2 speech production by L1 Spanish learners.

and realistic and reflects the occurrence of the various speech sounds in real-life speech. In addition, we analyzed an abundant amount of material per speaker to collect a sufficient number of exemplars for each speech sound.

Recordings were made at the university language laboratory. A complete recording, including tasks and responses, took approximately 30 minutes. Exams to test A1 proficiency level consisted of short tasks in which students had to give short answers taking into account the language use situation presented. Exams to test A2 and B1 proficiency had two parts. In part 1 students were asked to give short answers, whereas part 2 was aimed at collecting longer answers in situations in which students had to take the role of one of the interlocutors. The different tasks they had to accomplish varied from giving directions, or instructions, describing a problem and looking for solutions to telling about their plans and explaining the reasons why they had chosen to do that. Exams at the B2 proficiency level comprised three parts which were arranged from easier to more difficult and in all of them students had to take the role of one of the interlocutors. In every task they were asked to explain what they wanted and why, giving one or two reasons. Sometimes they had to describe a problem, to make different suggestions in order to solve it and to explain which decision they would make. For our study we only used those parts of the recording that contained the students' answers.

2.3.3 Data analysis

The students' responses were orthographically transcribed in Praat (Boersma & Weenink, 2010). For all transcriptions we used the SAMPA (Speech Assessment Methods Phonetic Alphabet; Wells, 2004) phonetic alphabet. Silences were removed from every recording, which resulted in three minutes of speech per recording, per speaker. A total of 69 minutes of running speech was available to be further analyzed. An automatic phonemic transcription

was generated using pronunciation variants from the lexicon of the Spoken Dutch Corpus (Oostdijk, 2002). Afterwards, manually corrected transcriptions that represent how the words were actually realized by our participants were generated. The transcriber, the first author, judged every sound and annotated deletions, substitutions and insertions. The transcriptions were corrected by using symbols that represent Dutch phonemes. One additional phoneme with allophonic status in Spanish (*/dʒ/*), was used.

Since it is well known that phonetic annotations tend to contain an element of subjectivity and that transcribers might be biased by their L1 (and in the present research, the transcriber was a native speaker of Spanish familiar with Spanish-accented Dutch) and their expectations of the results, we took measures to minimize these drawbacks. We decided to bring in a second annotator (a native speaker of Dutch not familiar with Spanish-accented Dutch) who tested the accuracy of the annotations by judging a sample of the transcriptions (10% of every recording). We then calculated the intertranscriber agreement by comparing the 10% of every recording that had been transcribed by the two annotators. Each pair of transcriptions was aligned automatically, and the discrepancies were located and tallied by an alignment programme. The results show to what degree the first and second annotator agreed in considering the target phonemes correct and in indicating mispronounced phonemes in the transcriptions. The degree of agreement between the first and the second annotator was high ($\kappa = .826$ for the vowels; $\kappa = .983$ for the consonants). Some vowels (*/ɛ/, /ə/, /œy/*) showed a mismatch percentage of more than 5%.

Further analyses revealed that the first annotator (a native speaker of Spanish) was more severe in judging whether target phonemes had been correctly realized than the second annotator (a native Dutch speaker), which is in line with results of previous studies on Dutch learners of English (Koet,

2007; Van den Doel, 2006), that showed that native speakers of Dutch were stricter in judging the English pronunciation of Dutch learners than English native speakers. To minimize possible biases the first annotator decided to go through every recording again and, if necessary, to make a second annotation. However, it turned out that it was not necessary to alter the transcriptions based on this second analysis, because the differences were negligible and thus would not add any relevant information to our study. Moreover, the experience of the first transcriber being a teacher skilled in phonology seemed to us solid enough to rely on her annotations above those of the second transcriber, a native speaker of Dutch who was not as trained in assessing pronunciation. As it is well-known, disagreement between transcribers is a common issue (Cucchiarini, 1996; Wester, Kessens, Cucchiarini, & Strik, 2001), especially when rating speech sounds produced by L2 students which are difficult to categorize as L2 phonemes, as is the case with /ɛ/, /ə/ and /œy/ (Van Doremalen, Cucchiarini, & Strik, 2013).

The annotations were used to generate confusion matrices comparing the automatically generated transcription (containing the target phonemes) with the manually corrected phonemic transcriptions (containing the realized phonemes). These matrices can be used not only to obtain scores on the overall performance of the participants, including their performance on consonants and vowels, but also to obtain detailed information about the realization of the target phonemes, the number of mispronunciations, and error percentages.

2.4 Results

We first calculated overall percentages of correctly pronounced segments per CEFR level. We distinguished three categories, as shown in Table 2.3 below. The category of Total correct includes, next to the vowels and consonants, the

insertions (inserted segments that are not part of the canonical pronunciation). The two other categories distinguish the consonants from the vowels.

Table 2.3 Means and standard deviations (SD) of proportions correct realizations per CEFR level, split out for total (including insertions), consonants and vowels.

Level	Total correct		Consonants correct		Vowels correct	
A1	.82	.04	.92	.02	.71	.06
A2	.92	.04	.98	.01	.86	.08
B1	.91	.03	.97	.01	.83	.06
B2	.93	.05	.98	.02	.86	.10
Overall	.90	.05	.97	.03	.82	.09

Table 2.3 indicates that the overall mean percentage vowel errors (18%) is relatively high in comparison to consonant errors (3%). One of every six vowels is erroneously realized, but the percentage may vary for the different vowels. There seems to be a gap between the A1 level and the rest (A2, B1, B2) in all three categories, but of course we have to keep in mind that this is not a longitudinal study and that the data refer to different learner groups. An ANOVA on the percentages total correct segments returned a significant result for proficiency level ($F = 6.785$, $p = .003$, partial eta squared = .517). The A1 informants have the lowest percentage correct and there is no overlap with the other three groups, all informants in the higher level groups (A2, B1, B2) have a higher score than the highest score in the A1 group. Within groups, between participants with the same level, the variation in percentage correct is fairly large, and there is a clear overlap among the levels A2, B1 and B2. Post hoc tests (Tukey HSD) for all three categories (total, consonants, vowels) revealed a distinction between A1 versus the higher levels (A2, B1, and B2), and no distinction between the three higher levels. This means that the decrease in pronunciation errors seems to taper off after the A1 level.

The percentage of correct realizations for consonants is high (97% overall) in comparison to vowels, which appear to be more problematic for Spanish learners of Dutch. The consonants show a significant effect of proficiency level as well ($F = 3.909$, $p = .025$, partial eta squared = .382).

Based on the overall percentage of correctly pronounced segments the question arises whether there is a relationship between the correct scores for consonants and vowels on the level of the participants. Making more vowel errors might be correlated to more consonant errors, but given the high correct scores for consonants one may assume that the acquisition of consonants is relatively effortless, independently of vowel pronunciation, which lags behind having its own pattern of acquisition. The correlation turns out to be high ($r = .803$, $p = .000$). The scattergram is given in Figure 2.3, with different symbols for the four proficiency levels.

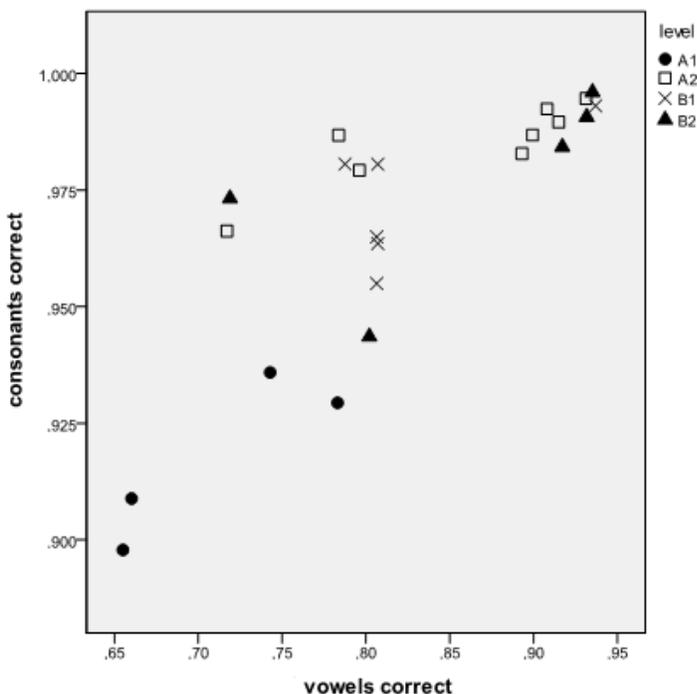


Figure 2.3 Scattergram of the correct proportions of vowel and consonant realizations for all participants (A1, A2, B1, B2).

As we can observe from the scattergram in Figure 2.3, there is a positive linear relationship between vowel and consonant scores, as indicated by the correlation of .803, but there is more to say, as there are many observations in the left upper half of the figure and none in the right lower half. This means that some participants had relatively fewer errors for consonants, compared to the others. This seems to indicate that vowel errors are not only more frequent, but also more persistent and variable.

In order to find the proper interpretation, we will need to have a look at the individual target phonemes. The distribution of errors may vary substantially for the different phonemes, whereas particular errors may be more typical of a specific proficiency level. Confusion matrices for target phonemes and

CEFR levels can provide us with the information required. We divided the data gained for vowels and consonants. While 8,447 (39.24%) of all target phonemes were vocalic phonemes, 13,075 (60.75%) represented consonantal phonemes. The percentages of vowel and consonant mispronunciations per target phoneme and CEFR level are specified in Table 2.4 and Table 2.5, as well as the frequency of occurrence of the target phoneme. The mispronunciations are split out on the phoneme level. We included insertions as well, the vowel insertions in Table 2.4, the consonant insertions in Table 2.5.

Below we will focus on those phonemes with a percentage error above 5%. Table 2.4 presents a full list of vowels displaying the pronunciation errors for 16 target vowels (the 15 unreduced vowels of Dutch plus the reduced vowel schwa). We give the relative frequency of occurrence of the errors (as reflected by percentages of erroneous pronunciations relative to the total number of occurrences of the phoneme).

At A1 level, the set of 15 unreduced target vowels contains six vowels with a mispronunciation percentage under 10% (/a:, ε, i, o:, ɔ, u/), eight above 10% (/ɑ, e:, I, y, ɣ, εi, œy, ø:/) and one vowel (/ɔu/) whose mispronunciation has not been observed. For the six vowels with a percentage of mispronunciation below 10% at A1 level, the error percentages were even lower for the three other proficiency levels. Table 2.4 shows that the phoneme /ɑ/ is often mispronounced as /a:/ (13.74%), but this problem is not found at the A2, B1 or B2 levels. The /e:/ is often realized as /ε/ (12.87%) or as /i/ (10.89%) at A1, but further problems above 10% of error percentage have not been found at other levels. The phoneme /I/ is frequently mispronounced as /i/ at A1 (52.31%), A2 (11.46%), B1 (16.61%) and B2 levels (38.18%). The /y/ is erroneously realized as /u/ at A1 (68.18%), A2 (30.43%), B1 (37.93) and B2

levels (12.50%). Substitutions of /y/ by /i/ have only been reported at B1 level (20.69%). The phoneme /ɣ/ was often mispronounced as /u/, having an error percentage of 46.15%. This outcome was similar to that of A2 (42.86%) and B1 (52.78%), whereas it was lower at B2 (22.22%). The phoneme /ø:/ was not reported at A1 level, but it was at B1 where it was mispronounced as /o:/ (50.00%). Although this percentage might seem high at first, it has to be interpreted in relation to a very low absolute number of only two occurrences. The diphthong /ɛi/ was substituted by /i/ at A1 level (10.71%). Substitutions of /œy/ by /ou/ were found at A1 (88.89%), A2 (33.33%), B1 (53.85%) and B2 (31.25%).

Incorrect realizations of the reduced vowel /ə/ are also shown in Table 2.4. Mispronunciations of the schwa as /ɛ/ were noticed at all levels (A1=45.22%, A2=33.53%, B1=47.83%, B2=23.17%).

Finally, we would like to comment on the vowel insertions we encountered in our study. Insertions of /ɛ/ (64.28%), /a:/ (14.28%) and /i/ (14.28%) were reported at A1. The noticeable insertions at A2 were caused by the phonemes /ɛ/ (36.36%), /ɪ/ (36.36%), and /ɛi/ (18.18%). At B1 level we also observe insertions of the phonemes /u/ (46.15%), /i/ (23.07%) and /ɛ/ (15.38%). Illegal insertions at B2 level were caused by the phonemes /a/, /ə/, /e:/, /i/ and /ɔ/, all of them with an error percentage of 11.11%.

Table 2.4 Frequency of vowel mispronunciations per target vowel category and CEFR level (A1, A2, B1, B2); T Ph=Target phoneme, N=Number of occurrences, %Error=Error percentage, Real=Realization, Del=Deletion, Ins=Insertion.

T Ph	Example	A1			A2			B1			B2		
		N	%Error	Real									
/a:/	naam	87	3.45%	/a/	100	-	-	182	1.10%	/a/	340	1.76%	/a/
/a/	kar	131	13.74%	/a:/	108	0.93%	/a:/	166	0.60%	/a:/	360	1.67%	/a:/
			1.53%	/e/		0.93%	/e:/						
			0.76%	Del		0.93%	/a/						
/e:/	teen	101	12.87%	/e/	97	5.15%	/e/	178	6.74%	/e/	315	6.67%	/e/
			10.89%	/i/		1.03%	/i/		1.12%	/i/		0.63%	/a/
			0.997%	/a/					0.56%	/e:/		0.32%	/i/
												0.32%	Del
/e/	pen	63	3.17%	/e:/	130	2.31%	/e:/	216	-	-	352	0.57%	/a/
			1.59%	/a/		1.54%	/a/					0.28%	/e:/
			1.59%	/i/									
/i/	tiem	75	4.00%	/i/	85	1.18%	/i/	140	-	-	152	0.66%	/e/
/i/	ik	130	52.31%	/i/	157	11.46%	/i/	271	16.61%	/i/	330	38.18%	/i/
			1.54%	/e/					0.74%	/e/			
									0.37%	Del			
/o:/	roos	61	9.84%	/a/	83	1.20%	/a/	101	3.96%	/o:/	192	8.33%	/a/
			1.64%	/a/					2.97%	/a/		1.56%	/o:/
/a/	rok	50	4.00%	/o:/	59	1.69%	/o:/	88	1.14%	/a/	186	2.15%	/o:/
			4.00%	/y/								0.54%	/a/
			2.00%	/u/								0.54%	Del
/u/	boek	37	2.70%	/a/	45	-	-	91	-	-	107	0.93%	/a/
/y/	duur	22	68.18%	/u/	23	30.43%	/u/	29	37.93%	/u/	40	12.50%	/u/
			4.55%	/y/		4.35%	/a/		0.74%	/e/			
						4.35%	/y/		6.90%	/e/			
/r/	buz	13	46.15%	/u/	14	42.86%	/u/	36	52.78%	/u/	54	22.22%	/u/
			7.69%	/o:/		7.14%	/i/					1.85%	/a/
/a/	lopen	272	45.22%	/e/	346	33.53%	/e/	577	47.83%	/e/	984	23.17%	/e/
			1.84%	/i/		2.89%	/i/		2.08%	/i/		0.51%	/i/
			1.10%	/e:/		1.16%	Del		0.52%	Del		0.41%	Del
			1.10%	Del		0.29%	/i/		0.17%	/e:/		0.30%	/i/
			0.74%	/i/		0.29%	/e:/					0.30%	/e:/
			0.37%	/a:/								0.20%	/e:/
			0.37%	/a/								0.10%	/a/
/o:/	neus	4	-	-	10	10.00%	/y/	2	50.00%	/o:/	12	8.33%	/o:/
						10.00%	/u/					8.33%	/u/
/e:/	trein	28	10.71%	/i/	97	2.06%	/e/	153	-	-	191	1.05%	/e/
/o:/	huus	9	88.89%	/o:/	9	33.33%	/o:/	13	53.85%	/o:/	48	31.25%	/o:/
/o:/	goud	9	-	-	18	-	-	21	-	-	47	-	-
Ins		14	64.28%	/e/	11	36.36%	/e/	13	46.15%	/u/	9	44.40%	/e/
			14.28%	/a:/		36.36%	/i/		23.07%	/i/		11.11%	/a/
			14.28%	/i/		18.18%	/e:/		15.38%	/e/		11.11%	/a/
			7.14%	/i/		9.09%	/i/		7.69%	/a/		11.11%	/e:/
									7.69%	/e:/		11.11%	/i/
												11.11%	/a/

Consonants display quite a different trend. The frequency of consonant mispronunciations and the category of insertions are presented in Table 2.5. The 21 consonant phonemes which are shown in Table 2.5 consist of 14 consonants with a mispronunciation percentage lower than 5% (/p, b, d, k, g, f, s, z, ʒ, m, n, ŋ, l, r/) and seven consonants with an error percentage above 5% (/t, v, x, h, ʃ, u, j/), which will be further discussed below.

Table 2.5 Frequency of consonant mispronunciations per target consonant category and CEFR level (A1, A2, B1, B2); T Ph=Target phoneme, N=Number of occurrences, %Error=Error percentage, Real=Realization, Del=Deletion, Ins=Insertion.

T Ph	Example	A1			A2			B1			B2		
		N	%Error	Real	N	%Error	Real	N	%Error	Real	N	%Error	Real
/p/	<i>pak</i>	73	1.37%	Del	65	-	-	94	-	-	145	-	-
	<i>boot</i>	50	-	-	53	-	-	66	-	-	131	-	-
	<i>rein</i>	162	23.46%	Del	284	4.23%	Del	467	7.92%	Del	718	4.60%	Del
			1.85%	/k/				0.43%	/n/		0.42%	/s/	
								0.21%	/d/		0.14%	/z/	
								0.21%	/ʃ/		0.14%	/d/	
/d/	<i>dutim</i>	116	-	-	176	1.14%	/n/	292	0.68%	/t/	481	0.21%	/t/
						0.57%	Del						
/k/	<i>kaat</i>	159	2.52%	Del	166	-	-	321	0.62%	Del	457	3.94%	/x/
			1.26%	/x/					0.31%	/x/		1.09%	Del
												0.44%	/f/
/g/	<i>goal</i>	1	-	-	0	-	-	2	-	-	2	-	-
/f/	<i>fijn</i>	47	2.13%	/p/	45	-	-	50	-	-	88	3.41%	del
			2.13%	Del									
/v/	<i>viz</i>	34	11.76%	/b/	77	2.60%	/b/	110	-	-	201	-	-
			2.94%	Del									
			5.88%	/o/									
/s/	<i>sein</i>	152	3.95%	Del	169	0.59%	/ʃ/	244	1.23%	Del	354	0.85%	/ʃ/
			3.29%	/ʃ/					1.23%	/ʃ/		0.56%	Del
									0.82%	/t/			
/z/	<i>zo</i>	25	4.00%	/s/	35	-	-	88	-	-	144	0.69%	Del
/x/	<i>toch</i>	59	10.17%	/t/	55	1.82%	Del	139	3.60%	Del	222	0.45%	/t/
			6.78%	Del					2.88%	/k/		0.45%	/s/
			1.69%	/g/								0.45%	/g/
			1.69%	/v/								0.45%	/t/
/h/	<i>hutz</i>	31	16.13%	/x/	54	3.70%	/x/	114	5.26%	/x/	160	1.25%	/x/
						1.85%	Del		0.88%	Del		1.25%	/k/
/y/	<i>bagage</i>	1	-	-	0	-	-	0	-	-	2	-	-
/ʃ/	<i>chique</i>	6	33.33%	/s/	7	-	-	0	-	-	5	-	-
			16.67%	/x/									
			16.67%	Del									
/m/	<i>met</i>	101	0.99%	/n/	110	-	-	207	-	-	312	0.64%	Del
												0.32%	/n/
/n/	<i>nee</i>	248	1.61%	Del	299	0.33%	Del	572	0.70%	Del	914	2.30%	Del
												0.11%	/x/
												0.11%	/k/
												0.11%	/t/
/ŋ/	<i>koning</i>	5	-	-	17	-	-	36	2.78%	/n/	53	1.89%	Del
									2.78%	Del			
/l/	<i>laat</i>	130	1.54%	Del	130	0.77%	Del	156	2.56%	Del	339	0.88%	Del
/r/	<i>rand</i>	149	0.67%	/n/	221	-	-	353	-	-	568	0.35%	Del
			0.67%	Del									
/v/	<i>wit</i>	49	30.61%	/b/	59	1.69%	/v/	127	9.45%	/b/	201	1.00%	/b/
												0.50%	/g/
/j/	<i>ja</i>	39	38.46%	/dʒ/	62	6.45%	/dʒ/	106	10.38%	/dʒ/	147	2.04%	/dʒ/
			5.13%	/z/		1.61%	Del						
Ins		31	48.38%	/d/	23	69.56%	/n/	40	80.00%	/n/	31	61.29%	/n/
			48.38%	/t/		21.73%	/d/		7.50%	/s/		19.35%	/s/
			3.22%	/x/		4.34%	/t/		5.00%	/x/		6.45%	/s/
						4.34%	/d/		2.50%	/g/		3.22%	/g/
										/v/		3.22%	/h/
												3.22%	/ʃ/
												3.22%	/t/

At A1 the phoneme /t/ is deleted in 23.46% of the cases in which it should have been pronounced. No significant number of deletions appeared at A2 and B2 levels, but there were deletions at B1 since the /t/ was deleted in 7.92% of the cases in which it occurred. The phoneme /v/ was substituted by /b/ at A1

level (11.76%). Problems at A2, B1 and B2 levels were not reported. The /x/ was mispronounced as /k/ (10.17%) or deleted (6.78%) at A1 level, whereas no serious problems were noticed at the other levels. The phoneme /h/ was mispronounced as /x/ at A1 (16.13%) and B1 levels (5.26%). Mispronunciations of the /ʃ/ as /s/ (33.33%) and /x/ (16.76%) or deletions (16.76%) were reported at A1 level and no pronunciation errors were found at A2, B1 or B2. The phoneme /v/ was substituted by /b/ at A1 (30.61%) and B1 levels (9.45%). Mispronunciations of /j/ in /dʒ/ were reported at A1 (38.46), A2 (6.45%) and B1 levels (10.38%).

To conclude, insertions of consonants were noticed at all levels, in particular insertions of /d/ and /n/, both with a percentage of 48.38% were found at A1. The consonantal phonemes /n/ (69.56%) and /d/ (21.73%) were also inserted at A2. Examples of insertions at B1 level were the phonemes /n/ (80.00%) and /s/ (7.50%). Finally, we encountered /n/ (61.29%), /s/ (19.35%) and /x/ (6.45%) as illegal insertions at B2 level.

2.5 Discussion

2.5.1 Vowels

The results of the present study are in line with those of previous research on pronunciation errors in Dutch L2 (Neri et al., 2006), which indicate that vowels are in general more problematic than consonants. The problems observed with Dutch vowels in Spanish learners are mostly related to vowel length in combination with the lax/tense distinction, but contrast in vowel height, rounding of front vowels and orthographic interference also appear to play a role.

Difficulties with vowel length are evident from the mispronunciations of the phonemes /a/ as /a:/: and /o:/: as /ɔ/, problems already shown in perceptual experiments reported by Escudero and Williams (2011, 2012). Spanish learners produce vowels that are closer to the Spanish phonemes /a/ and /o/. Orthographic interference may also cause mispronunciations of /a/ and /ɔ/, since in Spanish the graphemes <a> and <o> correspond to the Spanish phonemes /a/ and /o/ respectively. As shown by our results, these pronunciation problems seem to disappear as exposure to L2 increases.

Problems regarding contrast in vowel height are noticeable in pronunciation errors concerning the phoneme /e:/:, which was often substituted by /ɛ/ and, to a lesser extent, /i/. Mispronunciations of the phoneme /ɪ/ also point to problems concerning contrast in vowel height and may have their origin in the difficulty in producing a new vowel contrast or even perceiving it. According to Flege, Bohn and Jang (1997), adult learners will ultimately produce new L2 vowels more accurately than similar L2 vowels. This might explain the difficulties in realizing the phoneme /ɪ/, which is similar to /i/. Moreover, orthography may also play a role as the Dutch phoneme /ɪ/ is often represented by the grapheme <i> which in Spanish corresponds to the phoneme /i/.

Front rounded vowels as the /y/, /ʏ/, /ø:/: and /œy/: are also responsible for an important number of mispronunciations. The phonemes /y/, /ʏ/ and /ø:/: are frequently substituted by /u/. Such errors are probably due to orthographic interference, since in Spanish the grapheme <u> corresponds to the phoneme /u/. Problems concerning the pronunciation of front rounded vowels are equally evident from the mispronunciation of the phonemes /œy/: as /ɔu/ and,

in a few cases, /ø:/ as /o:/. The diphthong /œy/ is often realized in back position resulting in /ɔu/, which is very similar to the Spanish diphthong /au/.

Orthographic interference is also evident in frequent mispronunciations of schwa /ə/ as /ɛ/ or /e:/ when the phoneme is represented by the grapheme <e>. These mispronunciations are noticeable in prefixes such as *ge-* as in *gekocht*, ‘bought’, or *be-* as in *beschrijven*, ‘to describe’, or *ver-* as in *vertalen*, ‘to translate’. Also in the suffix *-en*, which indicates the infinitive, like in *kopen*, ‘to buy’. In Table 2.4 it is shown that /ə/ is often mispronounced as /ɪ/ or /i/, although to a much lesser extent than /ɛ/. This occurs in words with the suffixes *-ig* such as *aardig*, ‘nice’ or *-lijk* like *eindelijk*, ‘finally’, which despite their orthographic form need to be pronounced with the reduced vowel /ə/. This also applies to the schwa being mispronounced as /ɛi/ and as /ɪ/ or /i/ like in *gemakkelijk*, ‘easy’. This is probably related to orthographic interference from the grapheme <ij>, which corresponds to the phoneme /ɛi/, but in the suffix *-lijk* should be pronounced as /lək/.

Table 2.6 below presents a synopsis of the vowel confusions per CEFR level (A1, A2, B1, B2) and their possible sources. As we mentioned before, orthography is likely to play a role in all confusions.

Table 2.6 *Vowel confusions in A1, A2, B1 and B2, + = % error > 5, - = % error < 5, Error source L=Length, H=contrast in vowel Height, F=Front round.*

Confusion		Levels				Error source
		A1	A2	B1	B2	
/ɑ/	→ /a:/	+	-	-	-	L
/o/	→ /ɔ/	+	-	-	+	L
/e:/	→ /ɛ/	+	+	+	+	H
/ʌ/	→ /i/	+	+	+	+	H
/e:/	→ /i/	+	-	-	-	H
/œy	→ /ɔu/	+	+	+	+	F
/ø:/	→ /o:/	-	-	-	+	F
/y/	→ /u/	+	+	+	+	F
/ʏ/	→ /u/	+	+	+	+	F / H
/ø:/	→ /u/	-	+	-	+	F / H

In Table 2.6, we can observe mispronunciations concerning length (L, which combines with problems in the lax/tense distinction), contrast in vowel height (H) and front round (F). When we look at the results of the confusions, we see that the realizations (/a:, ɔ, ɛ, i, ɔu, o:, u/) are always close or similar to the five Spanish vowels (/a, o, e, i, u/) or the Spanish diphthong /au/. The error sources are good indicators of the kinds of problems, but when we turn the picture from Dutch to Spanish we see that the five Spanish vowels seem to function as attractors for the Dutch vowels, to a certain extent, putting them in a Spanish perspective or framework. The way the vowels pattern is summarized in Table 2.7, where we added the confusion patterns of the other Dutch vowels. A Dutch vowel is ordered under the Spanish vowel that has the highest matching scores.

Table 2.7 *The five Spanish vowels as attractors of 15 Dutch vowels; /ɔu/ not included, as no pronunciation errors were found.*

Spanish vowel attractor	Dutch vowels attracted
/a/	/a:/, /ɑ/
/e/	/e:/, /ɛ/, /ə/
/o/	/o:/, /ɔ/, /ø:/
/i/	/i/, /ɪ/, /ɛi/
/u/	/u /, /y/, /ʏ/, /œy/

Table 2.7 indicates the impact of the L1 phonological system on L2 vowel production. The feature organization of the Spanish phoneme inventory seems to be carried over into Dutch. According to the Speech Learning Model (SLM; Flege, 1995) the L1 and L2 phonetic subsystems of a bilingual share the same phonological space, which results in the interaction of the two phonetic subsystems (Flege, Schirru, & Mackay, 2003). However, the phonetic categories of the L1 are hypothesized to become more powerful attractors of L2 vowels (and consonants) due to maturational constraints, especially in initial stages of L2 learning (Parnell & Amerman, 1978; Walley & Flege, 2000). The participants involved in our study, Spanish speaking adult learners of Dutch, tend to fall back on their L1 vowels. They often produce L2 sounds using unmodified L1 phonetic segments, as displayed in Table 2.7.

The Dutch /ɪ/ is the lax counterpart of the Dutch tense vowel /e:/, but for the Spanish learners the lax counterpart is the /ɛ/ and the /ɪ/ is subsumed under the attractor /i/. This process is mirrored by the Dutch /ʏ/, the lax front round vowel, which is realized as the round back vowel /u/. These shifts might be strengthened by the influence of orthography since in Spanish the graphemes <i> and <u> correspond to the Spanish phonemes /i/ and /u/ respectively, whereas in Dutch they correspond to either /ɪ/ or /i/ and /ʏ/ or /y/.

The pronunciation errors observed are in line with cross-language speech perception models (Best, 1995; Escudero, 2005; Flege, 1995) which state that L2 learners will categorize and assimilate non-native sounds with regard to the phonetic categories of their native language and that the more similar the L2 sounds are to the phonemes of the native language, the more difficult it is to perceive and to produce them.

2.5.2 Consonants

The most frequent consonant errors are related to clusters and to single phonemes in word-initial and word-final position. Problems with clusters lead to insertions of /ε/ or /ə/ in the combinations /sp/ as in *Spanje*, ‘Spain’, /st/ as in *sterk*, ‘strong’ /sx/ as in *schoon*, ‘clean’, /sl/ as in *slim*, ‘smart’. The /h/ in initial position is often substituted by /x/ or deleted. The phonemes /v/ and, to a lesser extent, /v/ are frequently substituted by /b/ in word-initial position. These substitutions might be due to the Spanish phoneme /b/ functioning as attractor of the Dutch /v/ and /v/, a phenomenon which has been observed for the vowels (Parnell & Amerman, 1978; Walley & Flege, 2000). These mispronunciations might also be due to the fact that in Spanish the graphemes , <v> and, occasionally, <w> often correspond to the phoneme /b/. Another word-initial phoneme which is regularly mispronounced is the /j/, being realized as /dʒ/, allophone of the Spanish phoneme /k/ (Hualde, 2005). This pronunciation error may hint to interference from English. Consonant errors in word-medial position are evident from the mispronunciation of the /ʃ/ sound as in *rekenmachine*, ‘calculator’, which was frequently substituted either by /s/ or /x/ and, in a few cases, deleted. The Dutch velar fricative /x/ as in *mag*, ‘I may’ is commonly replaced by the velar plosive /k/, a mispronunciation that has been addressed in previous studies with Dutch L2

learners with different L1s (Neri et al., 2006; Strik, Truong, De Wet & Cucchiaroni, 2009). The phoneme /t/ in word-final position as in *moest*, ‘must, past tense’ is often deleted. Clusters within a word such as *buurtfeest*, ‘neighbourhood party’, or in word-final position as in *alstublieft*, ‘please’ and *markt*, ‘market’ were most likely responsible for most of the /t/ deletions, which is a common phenomenon among native Dutch speakers and is more likely when the preceding consonant is a fricative (Goeman, 2002).

Table 2.8 below presents a synopsis of the consonant confusions per CEFR level (A1, A2, B1, B2) and their sources.

Table 2.8 *Consonant confusions in A1, A2, B1 and B2, + = % error > 5, - = % error < 5, error source I¹=Interference from L1, I²=Interference from English, S=Syllable (CV), N=Novel phoneme.*

Confusion	Levels				Error source
	A1	A2	B1	B2	
/t/ → Del	+	-	-	-	I ¹ / S
/x/ → /k/	+	-	-	-	I ¹ /S
/h/ → /x/	+	-	+	-	N
/ʃ/ → /s/	+	-	-	-	N
/ʃ/ → /x/	+	-	-	-	N
/ʃ/ → Del	+	-	-	-	N
/j/ → /dʒ/	+	+	+	-	I ²
- → /n/	+	+	+	+	I ¹ / S
- → /d/	+	+	-	-	I ¹ / S
- → /s/	-	-	+	+	I ¹ / S
- → /x/	-	-	+	+	I ¹ / S
- → /g/	-	-	+	-	I ¹ / S

2.6 Conclusions and future perspectives

The aim of the present study was to gain insight into the pronunciation errors made by Spanish learners of Dutch L2. Our research has produced a detailed

overview of vowel and consonant errors from which we have tried to identify the most frequent errors and their possible sources. The resulting inventory can be used as a guideline in developing Computer Assisted Pronunciation Training (CAPT) systems that make use of Automatic Speech Recognition (ASR) to provide instantaneous and personalized feedback on Dutch L2 pronunciation.

Since our study was not longitudinal, the results for the various CEFR levels cannot be interpreted as progress of individual students, but they do indicate that some pronunciation errors were specific for the A1 level and disappeared at A2 level. Other mispronunciations remained at B1 and even at B2 levels and can be considered as examples of persistent errors. Not all participants within the same CEFR level performed equally on their speaking task, and some of them performed clearly worse in comparison to other peer-level students.

The mispronunciations observed in our study are in accordance with cross-language speech perception models which describe the interference from the L1 sound system on the acquisition of L2 speech sounds and specify that L2 sound contrasts that are mapped onto an L1 single category are the most difficult to discriminate and to learn (Best, 1995; Escudero, 2005; Flege, 1995).

Having made this consideration, we would like to draw the following conclusions from the results of our study. First, vowel errors are not only more frequent, but also more persistent and variable than consonant mispronunciations. Second, in producing Dutch vowels, Spanish learners appear to have problems with the contrast in vowel length and vowel height, and with front rounded vowels, as we already predicted. Problems with the abovementioned vowel features might be explained by the nature of their native language. Spanish does not have contrastive vowel length nor front

rounding. As to the vowel height, in Spanish there are two high vowels /i, u/, two mid vowels /e, o/, and one low vowel /a/ (Hualde, 2005). A rather simple height dimension compared to Dutch with close /i, y, u/, half close /ɪ, e:, ʏ, ø:, o:/, half open /ɛ, ɔ/ and open vowels /ɑ, a:/. (Booij, 1995). These factors may explain why Spanish learners of Dutch frequently fail in establishing new categories for phonic elements of the L2. Third, consonant mispronunciations are found in onset and coda clusters which, according to our predictions, are responsible for a considerable number of insertions, substitutions and deletions. Fourth, we have noticed that mispronunciations in both vowels and consonants might be strengthened by the influence of orthography.

Patterns of segmental confusion reported on in this study provide useful information to develop dedicated pronunciation training programs which focus on the specific problems we noticed. Further analysis will be required to gain more insight into the nature of these errors and the specific context in which they occur. However, on the basis of these results some suggestions can already be presented.

First, several errors appear to be caused by L1-L2 orthographic interference. It is even questionable whether such errors should be considered as real pronunciation errors. The type of training to repair those errors should in any case focus on the grapheme-phoneme relationships for the Dutch sounds in question, rather than on their specific articulatory or acoustic properties. It is to be expected that such orthography-related errors should be easier to correct than errors related to perception or production difficulties, which is partly supported by our data as many of these errors are less frequent at higher proficiency levels.

Second, the vowel data in Table 2.6 reveal difficulties with contrasts in length, height and front rounding. Minimal pair exercises focused on perception and production could be used to point out such contrasts, as these

appear to be particularly helpful and are appreciated by L2 learners (Neri, Cucchiarini & Strik, 2008).

Production training could include tasks with visual stimuli in addition to read aloud tasks to give the possibility of practicing speech production in a setting in which the influence of orthography on pronunciation is more limited (Erdener & Burnham, 2005; Silveira, 2007). In addition, articulatory information could also be provided, either pre-emptively or in the form of feedback, through graphical displays of the vocal tract, to direct the learner's attention to specific articulatory properties of Dutch phonemes, like for instance the difference between the velar fricative /x/ and the velar plosive /k/.

Furthermore, CALL-based interactive perception and production exercises for vowels like those in Wik, Hincks, and Hirschberg (2009), and Wik and Escribano (2009) might also be employed. The system developed by Wik and Escribano (2009) to support vowel acquisition in Swedish L2 is particularly innovative. Formants are tracked, a 3D ball moves over a vowel-chart canvas in real time and target spheres are placed at the target values of the vowels while the students' task is to get the target spheres. The system also employs a calibration technique that elicits cardinal vowels from the user and employs those to normalize the vowel-space canvas, thus catering for any user, regardless of vocal tract size.

Finally, in addition to perception and production tasks that focus on problematic sounds, it is also important to include exercises that range in complexity from separate words to sentences, paragraphs, and free speech (Moyer, 1999), which are more representative of the tasks L2 learners will be expected to perform in the L2 (Jones, 1997).

3 Spanish-accented Dutch vowel productions: Duration and spectral features

This chapter has been reformatted and slightly modified from:

Burgos, P., Jani, M., Van Hout, R., Cucchiarini, C., & Strik, H. Spanish-accented Dutch vowel productions: Duration and spectral features (submitted a).

3.1 Introduction

Adult learners have difficulties in acquiring the phonology of an additional, second language (L2) (Birdsong & Molis, 2001; Long, 1990), often because of native language (L1) interference (Cutler, 2012; Flege, Schirru, & MacKay, 2003). Several speech perception models, such as the Speech Learning Model (SLM; Flege, 1995), the Perceptual Assimilation Model (PAM; Best, 1995) and its extension the PAM-L2 (Best & Tyler, 2007), and the Second Language Linguistic Perception Model (L2LP; Escudero, 2005), have tried to predict and explain learners' difficulties in mastering an L2 phonological system in terms of the perceived similarity of segments in L1 and L2. While Flege's (1995) SLM focuses on learning and predicts constraints in the perception and production of L2 segments by L2 learners, Best's (1995) PAM predicts listeners' difficulties in the perceptual assimilability of non-native segmental contrasts to native categories. Its extension, the PAM-L2 (Best & Tyler, 2007), focuses on the learners' perception of L2 contrasts. In line with the SLM and the PAM, Escudero's (2005) L2LP proposes the Full Copying hypothesis which states that learners will initially perceive L2 segments as copies of their L1 native categories (see also Van Leussen & Escudero (2015) for a revision of the L2LP model).

The interaction between the L1 and L2 phonetic systems has been widely investigated, especially in relation to vowels (Baker & Trofimovich, 2005; Flege, Schirru, & MacKay, 2003; Iverson and Evans, 2007). Iverson and Evans (2007) found that the size and complexity of the L1 vowel inventory affect L2 vowel learning: native speakers of L1s with relatively small and simple vowel inventories (Spanish and French) achieved lower accuracy in recognizing English vowels than speakers of L1s with larger and more complex vowel systems (German and Norwegian).

The aim of the present study is to investigate Dutch vowel production accuracy by adult Spanish learners. More precisely, we will analyze – on the basis of acoustical measurements of duration and spectral features – whether the Dutch vowels produced by Spanish learners match those of native Dutch speakers, and whether these learner realizations are influenced by the properties of Spanish L1 vowels. This is particularly interesting given the notable differences between the Dutch and Spanish vowel systems.

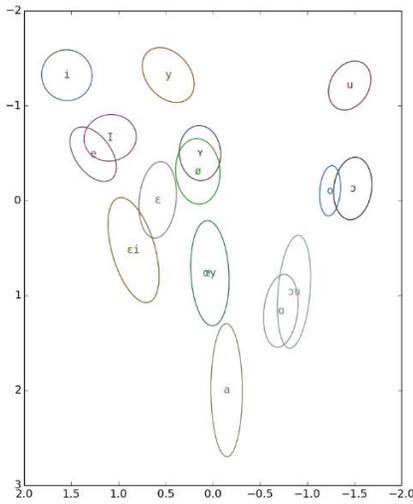


Figure 3.1 *F1* (*y*-axis) and *F2* (*x*-axis) (normalized) values for the 15 Dutch vowels, as realized by native Dutch speakers (measured at 50% of the vowel duration); the mean values are indicated by the vowel symbols; the values for Dutch vowels were drawn from Van der Harst (2011).

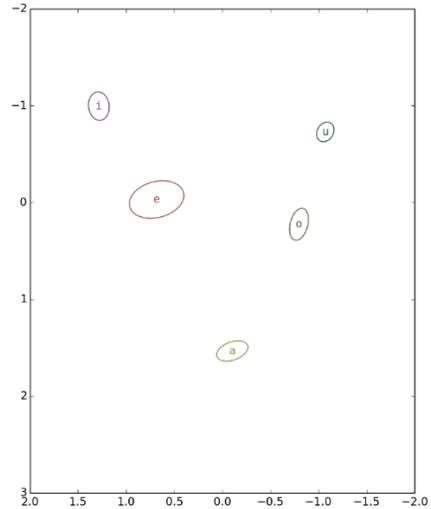


Figure 3.2 *F1* (*y*-axis) and *F2* (*x*-axis) (normalized) values for the five Spanish vowels, as realized by native Spanish speakers (measured at 50% of the vowel duration); the mean values are indicated by the vowel symbols; the values for Spanish vowels were drawn from Chládková, Escudero, and Boersma (2011).

Dutch has 15 full vowels (monophthongs: /i, y, u, ɪ, ʏ, ɔ, ε, a, a:/; long mid vowels: /e:, ø:, o:/; diphthongs: /ɛi, œy, ou/; see Figure 3.1), as well as the reduced vowel /ə/ (Adank, Van Hout, & Smits, 2004b; Booij, 1995), whereas Spanish has a five-vowel system (/a, e, i, o, u/; see Figure 3.2) (Hualde, 2005). Four features characterize differences between the two systems (see Table 3.1). Firstly, Spanish does not have phonemic vowel length (Hualde, 2005; McAllister, Flege, & Piske, 2002), whereas Dutch has a strict lax/tense distinction (lax vowels: /ɪ, ε, ɔ, ʏ, a/; tense vowels: /i, y, u, e:, ø:, o:, a:/), which

crosses the short/long distinction (short vowels: /i, y, u, ɪ, ʏ, ə, ε, ʌ/; long vowels: /a:,e:, ø:, o:, ei, œy, ɔu/) (Adank et al., 2004b). Dutch has four lax/tense vowel pairs, /ɪ/-/e:/, /ʏ/-/ø:/, /ə/-/o:/, /ʌ/-/a:/, whose lax vowels are short and whose tense vowels are long. Secondly, Spanish does not have front rounding, as all rounded vowels in Spanish are back vowels (/o, u/) (Hualde, 2005), while Dutch has four front rounded vowels (/y, y, ø:, œy/). The third feature pertains to diphthongs. Spanish does not have diphthongs at the phoneme level, that is, single phonemes defined by their trajectory between two vowel positions; instead it has 14 vowel combinations (Hualde, 2005). Dutch, on the other hand, does have diphthongs at the phoneme level, such as /ei, œy, ɔu/. The Dutch long mid vowels (/e:, ø:, o:/) are not considered to be full diphthongs, but are slightly diphthongized (cf. Adank et al., 2004b; Van der Harst, Van de Velde, & Van Hout, 2014). Finally, Spanish distinguishes three height values (high: /i, u/; mid: /e, o/; low: /a/) (Hualde, 2005), whereas Dutch is characterized by four (high: /i, y, u/; high mid: /ɪ, ʏ, e:, ø:, o:/; low mid: /ə, ε, ei, œy, ɔu /; low: /ʌ, a:/) (Booij, 1995).

Table 3.1 *Distinctive features of the Spanish and Dutch vowel systems.*

	Dutch	Spanish	Comment
Length (duration)	+	-	Dutch has a complex interaction between the distinctions of short/long and lax/tense
Front rounding	+	-	Both languages have back rounded vowels
Diphthong	+	-	Dutch has diphthongs at the phoneme level, whereas Spanish has 14 vowel combinations
Height	4 levels	3 levels	Dutch has an additional height value, as can be observed in the vowel space of Figure 3.1

Most studies on Spanish L1 Dutch L2 have focused on vowel perception. Researchers have investigated Dutch vowel perception by naïve Spanish listeners (Escudero & Williams, 2011), Spanish listeners with limited exposure to Dutch (Goudbeek, Cutler, & Smits, 2008; Escudero, 2015), and Spanish L2 learners (Escudero, Benders, & Lipski, 2009; Escudero & Wanrooij, 2010; Escudero & Williams, 2012; Escudero, Simon, & Mulak, 2014). Escudero and Williams (2011) found that naïve Spanish learners assimilated five Dutch monophthongs /i, u, ə, ε, a:/ and two Dutch long mid vowels /e:, o:/ primarily to a single Spanish vowel category /i, u, o, e, a/ (i.e., the five Spanish core vowels) or to a single Spanish vowel combination /ei, ou/. They categorized other vowel tokens of Dutch /ɪ, y, ʏ, ɑ, ø:/ in terms of two or more Spanish vowel categories, namely /i/ or /e/, /i/ or /u/, /e/ or /u/, /a/ or /o/ and /e/ or /ei/ or /eu/. Goudbeek et al. (2008) investigated the acquisition of novel Dutch phonetic categories /y, ʏ, ø:/, all three front and rounded, by Spanish listeners with limited Dutch exposure. In their experiment, distributional properties of the input (duration and vowel height) and availability of supervision (supervised *vs.* unsupervised learning) were varied across conditions. Their findings revealed that for the vowels investigated, Spanish listeners resorted primarily to F1 information (vowel height) to categorize the Dutch novel contrast /y/-/ø:/ (distinguished primarily by duration in native Dutch), and that this categorization was predominant when supervised learning was employed. Importantly, Goudbeek et al. (2008) concluded that Spanish learners, and L2 learners in general, find it extremely difficult to simultaneously use more than one cue to make a contrast (cf. Cutler, 2012). However, Spanish listeners do not solely resort to vowel height to categorize Dutch vowels. Escudero et al. (2009) showed that duration (instead of spectral cues) was the primary perceptual cue for Spanish learners

when categorizing the Dutch vowel contrast /ɑ/-/a:/ (based on both duration and vowel height in native Dutch). These findings confirm Goudbeek et al.'s (2008) results in showing that Spanish learners have difficulty simultaneously applying more than one cue to a contrast. Indeed, Escudero and Williams (2012) found – in a categorical discrimination task and a forced-choice identification task – that for Spanish learners the Dutch vowel contrasts /ɑ/-/a:/, followed by /I/-/i/, was the most difficult to discriminate. They suggest that this is because the two L2 phones in each pair are non-contrastive in Spanish as both resemble the Spanish /a/ and /i/ respectively.

Spanish orthography can also influence Dutch vowel perception. Spanish has a transparent orthography (i.e., phoneme-grapheme correspondence is straightforward), whereas Dutch has a deeper orthography (i.e., phoneme-grapheme correspondence is less clear-cut) (cf. Escudero & Wanrooij, 2010). Learning Dutch L2 vowels appears to be impeded when Dutch spelling conventions do not match phoneme-grapheme correspondence in Spanish (Escudero et al., 2014), especially in the case of the perceptually difficult contrasts /ɑ/-/a:/ and /I/-/i/ (Escudero, 2015). Besides, the use of digraphs in Dutch, such as <uu>, <aa>, <ee>, <oo> to represent the Dutch vowels /y, a:, e:, o:/, might induce lengthening of these vowels in Spanish learners who are unfamiliar with digraphs in L1 orthography (Escudero & Wanrooij, 2010; Escudero et al., 2014).

Few studies have addressed Spanish learners' speech production of Dutch (Burgos, Cucchiarini, Van Hout, & Strik, 2013, 2014a; Burgos, Jani, Cucchiarini, Van Hout, & Strik, 2014b). Research by Burgos et al. (2013, 2014a) showed that Spanish learners' vowel errors were more frequent and persistent than consonant mispronunciations. They had problems with contrasts in vowel length, vowel height, backness, and front rounding, and many of these problems were associated with L1 constraints. Annotations of

learner productions of the Dutch target vowels /ɑ, ɛ, ɪ, ɔ, ʏ/ were closer to the Spanish vowels /a, e, i, o, u/. Also, the Dutch long mid vowels /e:/ and /o:/ were often overdiphthongized as /ei/ and /ɔu/, resembling the Spanish vowel combinations /ei/ and /au/ respectively. Burgos et al. (2014b) acoustically analyzed Spanish learners' Dutch vowel productions in read speech. Acoustical analyses of three Dutch vowel contrasts, i.e., /ɑ/-/a:/ (both distinguished by place of articulation and duration), /ɪ/-/i/ (distinguished by place of articulation), and /ʏ/-/ø:/ (distinguished by duration) showed that learners did not employ duration and spectral properties in a native-like manner. Durational cues were primarily used to produce the /ɑ/-/a:/ and /ɪ/-/i/ contrasts.

The present investigation is a follow-up study to Burgos et al. (2014b), which investigated only six Dutch vowels; in the present study, we analyze Spanish learners' production of all 15 Dutch vowels. The aim is to investigate how native speakers of a straightforward five-vowel system, Spanish, produce L2 vowels from a more complex vowel inventory (i.e., Dutch). We will acoustically analyze speech data of a sample of adult learners with varying degrees of proficiency in Dutch L2 to obtain an overview of Spanish L1 Dutch L2 vowel pronunciations. An analysis of learner productions of all 15 Dutch vowels will enable us to understand how Spanish learners use specific dimensions or features to produce them. The central research question is: *Do the Dutch vowels produced by adult Spanish learners acoustically match those of native Dutch speakers?*

3.2 Five predictions

We address the central research question using Flege's (1995) SLM, which best suits our purposes since it focuses on learning and predicts constraints in the perception and production of L2 segments by L2 learners. Importantly, SLM explicitly addresses the question of creating *new* L2 categories, which is one of the foci of the present study. Flege's SLM predicts learners' difficulties in terms of an L1-L2 comparative approach based on the interaction of two mechanisms, i.e., equivalence classification and the formation of new categories. The mechanism of equivalence classification provokes that L2 learners erroneously interpret L2 segments as equivalents (i.e., *identical* or *similar*) to their own L1 categories. As a result, these L2 segments may differ considerably from native productions of the same speech segments. In the case of equivalence classification, the following two situations may arise: one L2 phoneme is matched to one L1 phoneme (e.g., there is a nearly direct match between the Dutch vowel /u/ and the Spanish /u/), or two distinct L2 segments fall into one single L1 category (e.g., the case of the Dutch vowels /ɑ/ and /ɑ:/ which are non-contrastive in Spanish as both resemble the Spanish /ɑ/). However, L2 segments that are sufficiently dissimilar from any L1 category (i.e., actually perceived as *new*) may evade the process of equivalence classification. In these cases, learners might be able to establish new phonetic categories for L2 segments. This may imply that in the case of Spanish L1-Dutch L2, the Dutch vowels /ɑ/ and /ɑ:/, which are non-contrastive in the L1 and can be considered *similar* to Spanish /ɑ/, will pose greater difficulties to Spanish learners than the *new* front rounded vowels /y/ and /ʏ/ which do not have a Spanish counterpart and are dissimilar from any native category.

On the basis of the acquisition studies on the Spanish L1-Dutch L2 pair mentioned earlier and Flege's (1995) SLM, we have formulated five

predictions to uncover the main pronunciation patterns of Dutch vowels produced by Spanish learners.

Prediction 1. Escudero and Williams (2011) found that Spanish listeners matched five Dutch peripheral vowels /i, ε, a:, ɔ, u/ to the Spanish core vowels /i, e, a, o, u/ respectively. The remaining Dutch peripheral monophthongs, /ɪ/ and /ʌ/, were ambiguously categorized as Spanish /i/ or /e/ and /a/ or /o/ respectively. They also found that the Dutch vowel contrasts /ʌ/-/a:/ and /ɪ/-/i/ were the most difficult to discriminate as both L2 segments in these contrasts resemble a single L1 category, namely, Spanish /a/ and /i/ respectively (Escudero & Williams, 2012). Burgos et al. (2013, 2014a) found the same confusion patterns. This leads us to predict that the five Spanish vowels will be matched to the five Dutch vowel categories /i, ε, a:, ɔ, u/ (prediction 1; equivalence classification according to Flege (1995); cf. the Full Copying hypothesis in Escudero (2005)¹).

Prediction 2. Flege (1995) suggests that the distance between a given category in the L2 and the closest native categories will influence the acquisition of *new* categories, making more distant categories easier to acquire. Dutch has four front rounded vowels /y, ʏ, ø:, œy/ which can all be classified as *new* for Spanish learners, as they do not have phonological counterparts in Spanish. These Dutch vowels are located in an empty area of the Spanish native vowel space, which might facilitate their acquisition by Spanish learners (Goudbeek et al., 2008, p. 123). However, previous research has shown that front rounded vowels can be particularly difficult for L2 learners whose native language does not have such vowels. Such L2 learners

¹ Escudero's Full Copy hypothesis states that L2 learners will initially perceive all L2 segments as exemplars of their native categories. This hypothesis resembles the mechanism of equivalence classification proposed by Flege (SLM; 1995).

appear to substitute the vowel /y/ with either /i/ or /u/, depending on their L1. This indicates that distinct languages divide the high vowel continuum differently (Rochet, 1995: 386). This, in turn, leads us to predict that adult Spanish learners will be capable of producing at least one *new* Dutch front rounded vowel category, somewhere in the higher spectral area of the vowel space, which has room to host new categories (prediction 2; new categories).

Prediction 3. As to the diphthongized Dutch long mid vowels (/e:/, ø:/, o:/) and the diphthongs (/ei, œy, œu/), Escudero and Williams (2011: 5) found that the first set were frequently categorized as the Spanish vowel combinations /ei, eu, ou/ respectively. Similarly, Burgos et al. (2014a) found that the long mid vowels /e:/ and /o:/ were often realized as Dutch /ei/ and /œu/, resembling the Spanish vowel combinations of /ei/ and /ou/ respectively. To avoid confusion with the Dutch long mid vowels, Spanish learners perhaps overdiphthongize the three Dutch diphthongs /ei, œy, œu/ by using the Spanish /a/ as the starting vowel. This is because the /a/ is the only Spanish vowel which could be used in combination with the Spanish high vowels /i/ and /u/. Our prediction is therefore that Spanish learners will overdiphthongize the Dutch long mid vowels and diphthongs, using Spanish vowel combinations as equivalents (prediction 3; equivalence classification, copy Spanish diphthongs; cf. the Full Copying hypothesis in Escudero (2005)).

Prediction 4. The fourth prediction involves duration, a highly relevant feature, like vowel height, in characterizing Dutch vowels (Adank et al., 2004b). Adult Spanish learners have difficulties in applying subtle spectral differences to make Dutch vowel distinctions (Burgos et al., 2013, 2014a, b). As a consequence, learners are likely to resort to duration, an accessible feature because Spanish has vowel combinations, the Spanish diphthongs, which are long. Bohn's (1995, p. 294–295) Desensitization Hypothesis posits

that learners whose linguistic experience has been insufficient to sensitize them to spectral differences that distinguish vowel contrasts in an L2 will resort to using duration differences to differentiate such contrasts. It is thus plausible that Spanish learners would resort to duration to expand the set of Dutch monophthongs, and especially to distinguish those vowels from pairs that 1) are based on subtle spectral differences, 2) do not exist in their native language and/or 3) cover areas of the acoustic vowel space in which a single L1 native category (or no L1 native category at all) is located (Bohn, 1995, p. 300). Such a strategy would imply that adult Spanish learners will employ duration to create new vowel categories in order to differentiate the Dutch vowel contrasts /ɑ/-/a:/, /ɪ/-/i/ and /ʏ/-/y/ (prediction 4; new categories using duration).

Prediction 5. The fifth and last prediction relates to the effect of L1 orthography on Dutch L2 vowel production (see also Escudero & Wanrooij, 2010; Escudero et al., 2014; Escudero, 2015). The prediction is that Dutch vowels, namely /y, a:, e:, o:/ which are represented by the digraphs <uu>, <aa>, <ee>, <oo>, will be produced with a longer duration by Spanish learners than by native Dutch speakers (prediction 5; orthography-vowel lengthening). If this prediction applies, it corroborates prediction 3.

The aforementioned predictions are presented in Table 3.2.

Table 3.2 Predictions based on Flege's (1995) Speech Learning Model and Escudero's (2005) Second Language Linguistic Perception Model.

	L1 Spanish	Spanish-Dutch vowels	Target L2 Dutch
Prediction 1 Equivalence classification (full copy Spanish monophthongs)	/i/	=====	/i/
	/e/	=====	/ɛ/
	/a/	=====	/a:/
	/o/	=====	/ɔ/
	/u/	=====	/u/
Prediction 2 New categories (front round)	-	/[ɣ~y]/	/ɪ/
	-		/y/
Prediction 3 Equivalence classification (copy Spanish diphthongs)	/ei/	=====	/e:/
	/eu/	→ /e[ɣ~y]/	/ø:/
	/ou/	=====	/o:/
	/ai/	=====	/ɛi/
	/au/	→ /a[ɣ~y]/	/œy/
	/au/	=====	/ou/
Prediction 4 New categories (using duration)	/a/	/a/-/a:/	/ɑ/-/a:/
	/i/	/i/-/i:/	/ɪ/-/i/
	-	/[ɣ~y]/-/[ɣ~y]:/	/ɪ/-/y/
Prediction 5 Orthography-vowel lengthening	-	/y:/	/y/ (<uu>)
	-	/a:/	/a:/ (<aa>)
	-	/e:/	/e:/ (<ee>)
	-	/o:/	/o:/ (<oo>)

3.3 Method

3.3.1 Spanish learners

A total of 28 highly educated adult Spanish learners of Dutch (9 males and 19 females) from Spain (21 learners; see Table 3.3) and Latin American countries (7 learners; Argentina, Dominican Republic, Guatemala, Mexico and Venezuela) took part in this experiment. Although we are aware of the phonetic differences among varieties of Spanish (Hualde, 2005), and their

possible influences on Dutch L2 perception (cf. Escudero and Williams (2012) in their study on Peruvian and Iberian Spanish learners of Dutch), we decided to pool all Spanish L1 Dutch L2 speech data in our study. For the focus of the present study, the perceptual differences reported by Escudero and Williams (2012) are negligible and do not bear out investigating these varieties of Spanish separately. In addition, phonetic differences between native speakers of Iberian and Latin American Spanish appear to be fewer when these speakers are highly educated (Navarro Tomás, 2004: 7), as the participants in the current study are.

All participants were living in the Netherlands at the time of this study, and reported being exposed to Dutch and using it daily. Their age of arrival (AoA) in the Netherlands varied between 19 and 42 years, and their age at the time of testing ranged between 20 and 52 years. Table 3.3 summarizes pertinent information about the participants per language proficiency level, in terms of the Common European Framework of Reference for Languages (CEFR; Council of Europe)². We investigated the Dutch vowel productions of learners of Dutch with four CEFR proficiency levels, namely, A1 (CES (= Cambridge English Scales) 100-119), A2 (CES 120-139), B1 (CES 140-159) and B2 (CES 160-180) (UCLES, 2015). All subjects had taken Dutch courses at some point during their stay in the Netherlands and were familiar with CEFR levels. They all rated their own proficiency level in Dutch, and in other foreign languages they spoke, using the CEFR Self-Assessment Grid (Council of Europe).

Our participants were multilingual in that they already spoke one or more languages before they started to learn Dutch. Six of them were

² The CEFR defines foreign language proficiency at six levels: A1, A2, B1, B2, C1, C2. These levels derive from three broad levels: A (basic user), B (independent user) and C (proficient foreign language user).

Spanish/Catalan bilinguals and had both languages as their L1. The foreign languages our subjects spoke were English ($n = 27$), French ($n = 9$), German ($n = 6$), Italian ($n = 4$), Portuguese ($n = 2$), Arabic ($n = 1$) and Euskara ($n = 1$). Dutch was an L3 or additional language (La) for all of them. As prior linguistic knowledge in multilinguals can be used during the acquisition of an La (De Angelis, 2007: 130), we looked at possible links between proficiency levels in other languages and proficiency level in Dutch. Participants' CEFR proficiency level in L2 English was variable and often high (0, $n = 1$; A1, $n = 1$; A2, $n = 0$; B1, $n = 3$; B2, $n = 8$; C1, $n = 10$; C2, $n = 5$). There was no significant correlation with their Dutch proficiency level and no significant correlations were found for the other languages either.

Table 3.3 presents percentage correct scores obtained in a study by Burgos, Sanders, Cucchiari, Van Hout, and Strik (2015). The vowel realizations studied in the current investigation were employed as speech stimuli in Burgos et al. (2015), in which non-expert native Dutch listeners, recruited through crowdsourcing, were asked to transcribe vowels produced by Spanish learners. "Correct" indicates that listeners transcribed the target vowels. The means and the ranges in Table 3.3 show that there are a large number of Dutch vowel errors in all four CEFR groups. An ANOVA on the percentage correct transcriptions yielded a significant effect for proficiency level ($F = 4.995$, $p = .008$, partial eta squared = .384). A Tukey HSD post-hoc analysis gives only a significant distinction between A1 vs. B2. The decrease in pronunciation errors seems to taper off slowly after the A1 level. The large overlap in correct scores between the proficiency levels shows that phonology acquisition does not always progress along with foreign language proficiency. These findings are in line with earlier findings in Burgos et al. (2014a), which reveal that adult Spanish learners, irrespective of their CEFR proficiency level in Dutch (B2 and lower), continue to have severe problems with the acquisition of

Dutch sounds, especially with respect to vowels. There is a significant correlation between length of residence (LoR) and percentage correct transcriptions ($r(28) = .387$, p (two-tailed) = .04), but this effect runs largely parallel to the CEFR level (ANOVA with LoR as the dependent variable returns $F(3,24) = 9.849$, $p = .000$, partial eta squared = .552). Post-hoc analysis shows a clear distinction in LoR between the A and B levels.

Table 3.3 *Scores for % Correct vowels, range of % Correct vowels, LoR, and number of participants per Spanish language variety per CEFR proficiency level.*

CEFR ^a	%Correct ^b	Range		LoR ^d	L1 ^e	
		%Correct ^c			IS	LAS
A1	58.5 (12.3)	37.9–79.8		2.0 (3.1)	8	2
A2	68.2 (8.0)	55.1–74.3		0.9 (0.4)	6	1
B1	68.4 (10.6)	57.7–79.7		8.0 (3.9)	2	2
B2	78.8 (10.5)	64.4–90.6		10.1 (5.9)	5	2
Overall	67.4 (12.9)	37.9–90.6		4.6 (5.3)	21	7
Mean (SD)						

a Language proficiency level according to the CEFR.

b Percentage correct Spanish L1 Dutch L2 vowel realizations, as transcribed by non-expert native Dutch listeners (Burgos et al., 2015).

c Range percentage correct Spanish L1 Dutch L2 vowel realizations, as transcribed by non-expert native Dutch listeners (Burgos et al., 2015).

d Length of residence in the Netherlands, in years.

e Spanish language varieties spoken by the participants, i.e., Iberian Spanish (IS) and Latin American Spanish (LAS).

As already mentioned, we pooled all Spanish L1 Dutch L2 speech data, to establish a cross-section of pronunciation diversity and variability. Our

motivation for this decision is supported by empirical studies that have provided evidence that factors such as a high overall proficiency level in the L2, a relatively long length of residence in the host country or substantial L2 use on a daily basis do not guarantee success in achieving a native-like pronunciation (cf. Flege, Frieda, & Nozawa, 1997; Munro, 1993; Yeni-Komshian, Flege, & Liu, 2000). These findings seem to suggest that phonology acquisition does not always reflect the level of foreign language proficiency, an assumption that allows us to pool all our speech data, as all our learners are below native-like level (C1 and C2) and none of them has a perfect pronunciation score.

3.3.2 Speech material

As mentioned, we used part of an existing corpus of Spanish L1 Dutch L2 (Burgos et al., 2014b) containing systematic productions of Dutch speech sounds that are problematic for Spanish learners. The part we analyzed consists of speech material produced by Spanish learners who read Dutch monosyllabic words from a computer screen. Read speech is often used in studies on L2 speech production (Chládková et al., 2011; Moyer, 1999), and especially on learners' vowel realizations. Obtaining controlled speech material enables us to investigate learner productions which systematically contain all target phones (see Burgos et al., 2014b). The word list reading task used to elicit data was previously used in Van der Harst (2011) and Van der Harst et al. (2014). It contained a total of 278 monosyllabic and disyllabic words representing all Dutch vowels in different contexts.

From these 278 words we selected a set of 29 monosyllabic Dutch words per speaker. The words contained all the Dutch vowels in stressed position followed either by /s/ or /t/, as it is known that the change in vowel quality is maximally reduced in alveolar contexts (Van der Harst, 2011: 146; Van der

Harst et al., 2014: 254). Table 3.4 provides an overview of the 15 Dutch vowels and their corresponding orthographic and phonological representations. No example of the vowel /y/ followed by /s/ was included, as this combination does not occur in Dutch monosyllabic words, except in proper names.

Table 3.4 Selected *-s* and *-t* words as speech stimuli (Van der Harst, 2011); Phon = phonological representation (in IPA), Orth = orthographic representation.

Vowel	Phon	Orth	s-word	t-word		
Monophthongs	/i/	<ie>	<i>kies</i>	/kis/	<i>riet</i>	/rit/
	/y/	<uu>	-	/y/	<i>fiut</i>	/fyt/
	/u/	<oe>	<i>poes</i>	/pus/	<i>voet</i>	/vut/
	/ɪ/	<i>	<i>vis</i>	/vɪs/	<i>fit</i>	/fɪt/
	/ʏ/	<u>	<i>zus</i>	/zʏs/	<i>put</i>	/pʏt/
	/ɔ/	<o>	<i>vos</i>	/vɔs/	<i>vlot</i>	/vlot/
	/ɛ/	<e>	<i>zes</i>	/zɛs/	<i>vet</i>	/vɛt/
	/ɑ/	<a>	<i>gas</i>	/xɑs/	<i>rat</i>	/rat/
	/aː/	<aa>	<i>aas</i>	/aːs/	<i>staat</i>	/staːt/
Long mid vowels	/eː/	<ee>	<i>mees</i>	/meːs/	<i>beet</i>	/beːt/
	/øː/	<eu>	<i>neus</i>	/nøːs/	<i>neut</i>	/nøːt/
	/oː/	<oo>	<i>boos</i>	/boːs/	<i>boot</i>	/boːt/
Diphthongs	/ɛi/	<ij>	<i>ijs</i>	/ɛis/	<i>spijt</i>	/spɛit/
	/œy/	<ui>	<i>huis</i>	/hœys/	<i>fluit</i>	/flœyt/
	/ɔu/	<ou>	<i>kous</i>	/kɔus/	<i>fout</i>	/fɔut/

The Dutch vowels realized by 20 native speakers of Standard Dutch (10 males and 10 females), which are presented in Table 3.4, were used in Van der Harst (2011) and Van der Harst et al. (2014) to describe the Dutch vowel system (see Figure 3.1).

A data set of read speech in native Spanish, previously used in Chládková et al. (2011), was also employed in the current investigation. The speech data consisted of “read words and sentences that were presented in Spanish orthography on a computer screen” (Chládková et al., 2011: 418). The Spanish vowels produced by 40 Spanish subjects (20 males and 20 females) are displayed in Figure 3.2, and were used as a reference for the Spanish learners’ Dutch vowel realizations (see Figure 3.5).

The methodology of defining vowel spaces we used is common in establishing vowel spaces for native speakers of specific languages. Examples were presented in Figure 3.1 and Figure 3.2 for Dutch and Spanish. Native vowel spaces are commonly based on read speech, although new computer packages make it possible to analyze large samples of spontaneous speech (cf. Rosenfelder, Fruehwald, Evanini, Seyfarth, Gorman, Prichard, & Yuan, 2014). Such a package is not yet available for Dutch.

3.3.3 Procedure

Recordings were made at the Linguistics Department of the Radboud University Nijmegen or at the participants' home. A headset (Logitech, USB entry DZL-A-0001 4-B) and laptop (ACER AMD Quad-core Processor A6-3400M with Turbo CORE Technology up to 2.30 GHz) were used for the recordings, which were all made in a quiet room. The data was recorded with a sampling frequency of 16 kHz. The words to be read out were presented on a computer screen, one by one, with three seconds between words. Each word was produced by each speaker only once. Prior to the reading task, subjects were asked to read five Dutch words on the computer screen to familiarize them with the task. Of the 812 recorded words (29 target words x 28 speakers) six (from six different subjects) were excluded due to erroneous recording.

This resulted in a total of 806 word recordings which were annotated, segmented and analyzed.

3.3.4 Analysis of the speech recordings

The words were orthographically transcribed in Praat (Boersma & Weenink, 2010). The vowels were subsequently segmented following the procedures described in Van Son, Binnenpoorte, Van den Heuvel, and Pols (2001), and Van der Harst (2011). We looked at information from the waveform, spectrogram, formant tracks and auditory cues to determine the beginning and end of each vowel. Segmentation was done by an experienced transcriber, the first author, and was then checked by a native Dutch phonetician.

3.3.5 Acoustical analyses

After vowel segmentation was completed, we conducted analyses of the durational and spectral characteristics of the Dutch L2 vowel realizations. To determine whether the Dutch vowels produced by the Spanish learners approached those produced by native speakers, we compared the results of the acoustical analyses with those obtained by Van der Harst (2011) and Van der Harst et al. (2014) for native Dutch vowels. Both acoustical analyses, i.e., for the Spanish learners and for the native Dutch speakers, were carried out on three groups of vowels: nine monophthongs, three long mid vowels, and three diphthongs (cf. Adank et al., 2004b; Van der Harst, 2011; Van der Harst et al., 2014). We then conducted formant extraction for every vowel. Acoustical analyses were carried out automatically to obtain measurements of duration, and first and second formants (F1 and F2). It should be noted that measurements of third formant (F3) were not analyzed in the current study, as previous research showed that F3 is not necessary for the identification of front rounded vowels in Dutch, and that using F1 and F2 only is sufficient to

identify these vowels (Adank, 2003; Cohen, Slis, & 't Hart, 1963,1967; Van der Harst, 2011). The first two formants were measured at three equidistant points (i.e., at 25%, 50% and 75% of the vowel duration). This information helps to determine if and how diphthongization is realized when producing all mid vowels and diphthongs by the Spanish learners in comparison to the native speakers, as mid vowels and diphthongs are long and show a milder or stronger degree of diphthongization in Dutch (Adank et al., 2004b; Van der Harst et al., 2014).

All measurements were automatically extracted using an LPC (Linear Predictive Coding) analysis in Praat (Boersma & Weenink, 2010). First, as an initial approximation, we used Praat to search for five formants in the range from 50 Hz to 5000 Hz for male speakers and 50 Hz to 5500 Hz for female speakers. Next, every vowel token was assigned a specific number of coefficients, i.e., four, five or six, based on information from the waveform, spectrogram and formant tracks of the speech signal. Subsequently, an LPC script based on the chosen number of coefficients was run in order to extract F1 and F2 values. The same procedure was repeated for the measurement extractions of duration. All measurements were manually checked by the first author and, where errors were found, these were corrected. Subsequently, an additional check for outliers was carried out following the procedure employed in Van der Harst (2011) and Van der Harst et al. (2014). Any outliers were carefully checked at 25%, 50% and 75% by the same native Dutch phonetician mentioned earlier, and corrected where necessary. Subsequently, Lobanov's (1971) z-score transformation was employed to normalize all vowel measurements in order to neutralize the formant frequency variations resulting from anatomic differences among informants (cf. Adank et al., 2004a; Van der Harst et al., 2014).

Vowel ellipses were computed for all vowel realizations. The ellipses computed for all 15 Dutch vowels (Figure 3.1) and for the Spanish vowels (Figures 3.2 and 3.5) relate to formant frequencies at 50% of the vowel duration. The ellipses computed for the Dutch monophthongs (Figures 3.4 and 3.5) are also computed at 50% of the vowel duration. The Dutch long mid vowels (Figures 3.6 and 3.7) and the diphthongs (Figures 3.8 and 3.9) were measured at three time points, i.e., at 25%, 50%, and 75% of the vowel duration, as explained in Van der Harst et al. (2014). The arrows show the 25%→50%→75% direction.

3.4 Results

The results of the acoustical analyses are presented in two subsections. In the first, we compare the duration of the native and the learner vowel realizations. In the second subsection, we consider the spectral features of the native and learner vowel realizations. First, the spectral features of the native and the learner vowel productions are studied separately. Subsequently, a comparison between the native and the learner data is made to determine whether Spanish learners produce Dutch vowels in a native-like manner and to identify possible confusions.

3.4.1 Duration

Figure 3.3 and Table 3.5 present the durations in milliseconds (ms) for all 15 Dutch vowels, for both the native Dutch speakers and the Spanish learners. The nine Dutch monophthongs (/i, y, u, l, ʏ, ɔ, ε, a, a:/) consist of eight short vowels and the long vowel /a:/. The eight monophthongs, which are phonetically short, as realized by native Dutch speakers, have an average duration of max. 110 ms, as can be inferred from Table 3.5. The vowel (/a:/), the long mid vowels (/e:, ø:, o:/) and the diphthongs (/ei, œy, ou/) are

phonetically long vowels in Dutch and their average duration, as produced by native speakers, is higher than 191 ms (see Table 3.5). When looking at the vowels produced by Spanish learners, we notice that their average values are consistently higher and much more variable (for all 15 vowels) than the average native values. Our findings show that Spanish learners make temporal distinctions between short and long vowels, although the way in which they employ duration is not native-like. A deviant case is the Dutch short vowel /y/, which is produced by the Spanish learners with a duration typical of long vowels, as displayed in Figure 3.3 and Table 3.5 (see also prediction 5).

Our outcomes indicate that the Dutch vowels produced by Spanish learners are systematically longer. Do these durational differences become a problem in perception? According to Nootboom and Doodeman (1980: 285), producing a speech segment with duration deviance between 25-100 ms “would possibly drastically upset the perception of the temporal structure of a speech utterance.” Table 3.5 shows that the average error of the learner durations for the vowels /ɔu, ɛ, l/, as compared to the native durations, is not higher than 25 ms, indicating that these vowels are produced within the range of perceptual tolerance. On the other hand, the learner durations of the /ɛi, ʏ, œy, a:, ɔ, ø:, o:, ɑ, u, i, e:, y/ (ordered from longer to extremely longer durations) are within the noticeable 25-100 ms range. Perhaps they will be perceived as longer, but none of the durations switches short vowels to the category of long vowels. The only exception is the /y/, with an average duration deviance of 120 ms.

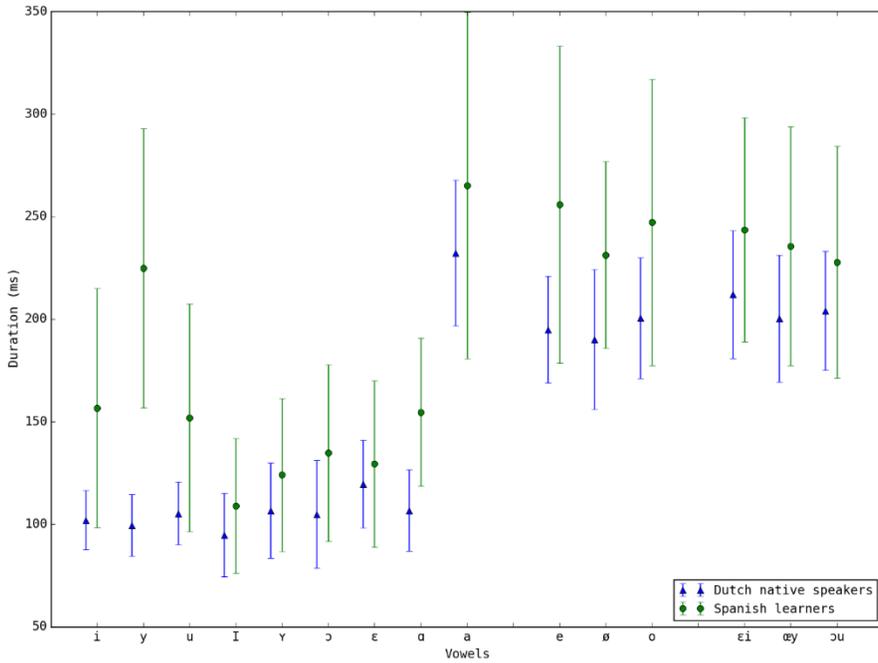


Figure 3.3 *Durational (raw) values for all 15 Dutch vowels, as realized by native Dutch speakers (triangles) and Spanish learners (circles). The values for Dutch vowels were drawn from Van der Harst (2011).*

Table 3.5 *Mean vowel durations (in ms) of Dutch short and long vowels produced by native Dutch speakers (DNS; Van der Harst, 2011) and Spanish learners (SL); %IDur=percentage of increase in vowel duration by Spanish learners.*

	Short vowels								Long vowels						
	Monophthongs								Long mid vowels				Diphthongs		
	/i/	/y/	/u/	/I/	/Y/	/ɔ/	/ɛ/	/a/	/aɪ/	/eɪ/	/øɪ/	/oɪ/	/ɛi/	/øy/	/ɔu/
DNS	97	105	99	88	97	97	110	104	235	192	191	198	218	207	214
SL	157	225	152	109	124	135	129	155	265	256	231	247	243	235	228
% IDur	61.8	114.2	53.5	23.8	27.8	39.1	17.2	49.0	12.7	33.3	20.9	24.7	11.4	13.5	6.5

3.4.2 Spectral features

In this section, we present the F1 and F2 values for all 15 Dutch vowels, as realized by native Dutch speakers and by Spanish learners. We first present the graphs, which display the ellipses representing the native and the learner vowel realizations, for the three groups of Dutch vowels separately, the monophthongs, the long mid vowels, and the diphthongs.

We apply multivariate analysis of variances (MANOVAs), first of all to the F1 and F2 spectral values as the dependent variables. Given the duration variation in the monophthongs (see e.g., the longer durations of the /y/ for the Spanish learners), we supplemented this analysis by including duration as well, to investigate whether the results improve or not. Duration was not included in analyzing the long mid vowels and diphthongs because they are long vowels (see Figure 3.3 and Table 3.5); their distinctive character is their spectral trajectory. To include this information, MANOVAs were performed on the F1 and F2 spectral values as dependent variables on three time points (25%, 50%, 75% of the vowel duration).

3.4.2.1 Monophthongs

The ellipses representing the realizations of the Dutch monophthongs (/i, y, u, l, ʏ, ɔ, ε, α, a:/) by native Dutch speakers and Spanish learners are displayed in Figure 3.4 and Figure 3.5 respectively. All ellipses in Figure 3.4 and Figure 3.5 are computed on the F1 and F2 values measured at 50% of the vowel duration.

The ellipses in Figure 3.4 show that most of the native vowel realizations occupy small and distinct areas in the vowel space. A MANOVA was conducted on all 36 pairs of the nine Dutch monophthongs (36), as realized by native Dutch speakers, to test how different the native vowels are. We used Pillai's trace, a multivariate measure of the proportion of shared variance

between two measurements, vowels in our case. A value of 1 means perfectly different or separate (perfect split), a value of 0 no difference at all (complete merger). All native vowel pairs of the Dutch monophthongs turn out to be significantly different in their F1 and F2 values, Pillai's trace always being high. With F1 and F2 as predictors, all native Dutch vowel pairs have values of above .800, except for /ɪ/-i/ (.704), /ɛ/-ɪ/ (.788) and /ɑ/-a:/ (.607). Adding duration hardly improves the outcomes, as expected, since monophthongs are phonetically short in native Dutch, with the exception of the /a:/. Increases in Pillai's trace are always less than .05, except for the /ɑ/-a:/ distinction, which increases to .816 when duration is added as a predictor. According to these outcomes, vowel distinctions are robust in native Dutch, the /ɪ/-i/ distinction being the most vulnerable one.

The ellipses representing the realizations of the nine Dutch monophthongs by Spanish learners displayed in Figure 3.5 are larger than those of native Dutch speakers in Figure 3.4. The ellipses displaying the learner realizations of the front rounded vowels /ʏ/ and /y/ occupy large portions of the vowel space, while the ellipses of other vowels, such as the /ɛ/ and /a:/, exhibit more restricted spatial areas. The ellipses displaying the learner realizations in Figure 3.5 show that there are six clear spectral areas. The presence of five spectral areas, which also coincide with the five Spanish vowels, are in accordance with prediction 1. The occurrence of a sixth spectral area meets prediction 2. Moreover, the learner vowel ellipses overlap each other. When looking at the ellipses displaying the learner realizations in Figure 3.5, we observe four non-overlapping areas, i.e., 1) /ɑ/ and /a:/, 2) /ɛ/, 3) /ɪ/ and /i/, and 4) /ɔ/, /u/, and /ʏ/ and /y/. The three areas /ɑ/ and /a:/, /ɛ/, and /ɪ/ and /i/ overlap with the ellipses representing the realizations of the Spanish vowels /a/, /e/ and /i/ respectively. The area displaying the ellipses of the vowels /ɔ/,

/u/, /ʏ/ and /y/ (see right upper part of Figure 3.5) contains three subsets (/ɔ/, /u/, and /ʏ/ and /y/). The ellipses representing the realizations of the Dutch /ɔ/ and /u/, two subsets which slightly overlap with each other, also overlap with the ellipses displaying the realizations of the Spanish /o/ and /u/ respectively (see prediction 1). The ellipses representing the vowel productions of the Dutch /ʏ/ and /y/ overlap considerably and constitute one subset ($[\text{ʏ}\approx\text{y}]$), in line with prediction 2. This subset also overlaps with the ellipsis representing the learner realizations of the Dutch /u/.

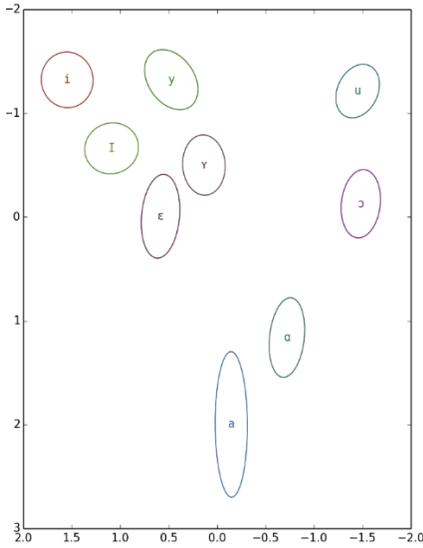


Figure 3.4 *F1* (*y*-axis) and *F2* (*x*-axis) (normalized) values for nine Dutch monophthongs, as realized by native Dutch speakers (50%); the mean values are indicated by the vowel symbols; the values for Dutch vowels were drawn from Van der Harst (2011).

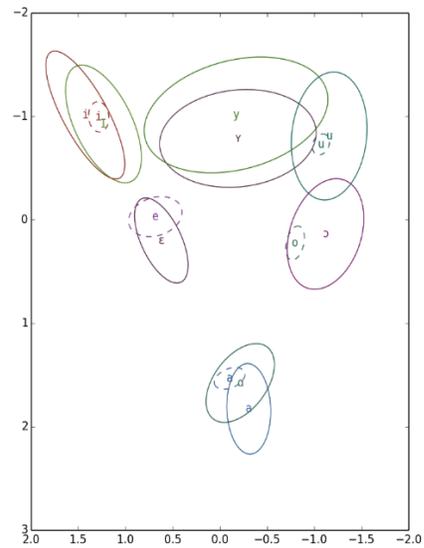


Figure 3.5 *F1* (*y*-axis) and *F2* (*x*-axis) (normalized) values for nine Dutch monophthongs, as realized by Spanish learners (50%); the mean values are indicated by the vowel symbols; the values for Spanish vowels (dashed lines) were drawn from Chládková et al. (2011).

To investigate the learner vowel distinctions of the nine Dutch monophthongs, we computed Pillai's trace values for all pairs using their *F1* and *F2* spectral values. The outcomes are given in Table 3.6.

Table 3.6 *Explained variance (multivariate, Pillai's trace) of the pairwise combinations of the nine Dutch monophthongs, as realized by Spanish learners, with F1 (50%) and F2 (50%) as predictors; values < .500 in grey, values > .500 and < .800 in light grey; 0=complete merger, 1=perfect split. All outcomes are statistically significant (alpha = .05).*

Vowel	i	y	u	ɪ	ʏ	ɔ	ɛ	ɑ	a:
i	-	.734	.949	.138	.349	.937	.626	.913	.923
y	-	-	.412	.672	.130	.747	.844	.936	.887
u	-	-	-	.936	.400	.439	.932	.894	.873
ɪ	-	-	-	-	.593	.930	.689	.923	.926
ʏ	-	-	-	-	-	.627	.690	.892	.859
ɔ	-	-	-	-	-	-	.902	.801	.772
ɛ	-	-	-	-	-	-	-	.838	.867
ɑ	-	-	-	-	-	-	-	-	.118
a:	-	-	-	-	-	-	-	-	-

The outcomes in Table 3.6 confirm the configurations in Figure 3.5. Three vowel pairs have very low values, i.e., /ɑ/-/a:/ (.118), /y/-/y/ (.130), and /ɪ/-/i/ (.138), indicating a low degree of separation. Most pairs (28 out of 36) have a value of above .607, the lowest boundary, with F1 and F2 as predictors, found for native Dutch (for the /ɑ/-/a:/ pair). However, the lower outcomes overall in Table 3.6 indicate that Spanish learners have difficulties when realizing the nine Dutch monophthongs.

Better results when adding duration as a dependent variable would indicate that Spanish learners use temporal cues to make vowel distinctions, as hypothesized in prediction 4. The results show that increases in explained variance by including duration are less than .05 in the majority of vowel pairs (29 out of 36). The learner vowel realizations which benefit the most from

adding duration as predictor are /i/, /y/ and /a:/, particularly in the case of the vowel pairs /l/-/i/ (F1 and F2 only .138; increasing to .677 after including duration), /ʏ/-/y/ (.130 → .495) and /ɑ/-/a:/ (.118 → .437). These results are in accordance with prediction 4. It should be noted that the results of the statistical analyses of the native data revealed that the /ɑ/-/a:/ is the only pair that benefits from adding duration, as distinctions between the native vowels in all other pairs of monophthongs are primarily based on spectral properties. Our results indicate that Spanish learners employ durational cues to distinguish the pairs /l/-/i/ and /ʏ/-/y/ as well, whereas these vowel contrasts are based on spectral cues in native Dutch.

3.4.2.2 Long mid vowels

The ellipses representing the realizations of the Dutch long mid vowels (/e:, ø:, o:/) by native Dutch speakers and by Spanish learners are presented in Figure 3.6 and Figure 3.7 respectively. The three ellipses displayed for each long mid vowel represent three points of measurements (25%, 50%, 75%).

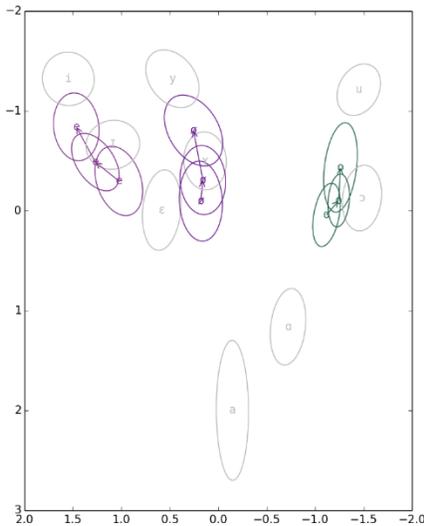


Figure 3.6 *F1 (y-axis) and F2 (x-axis) (normalized) values for three Dutch long mid vowels, as realized by native Dutch speakers (25%, 50%, 75%); the position of the nine Dutch monophthongs produced by native Dutch speakers is indicated in gray; the mean values are indicated by the vowel symbol; the values for Dutch vowels were drawn from Van der Harst (2011).*

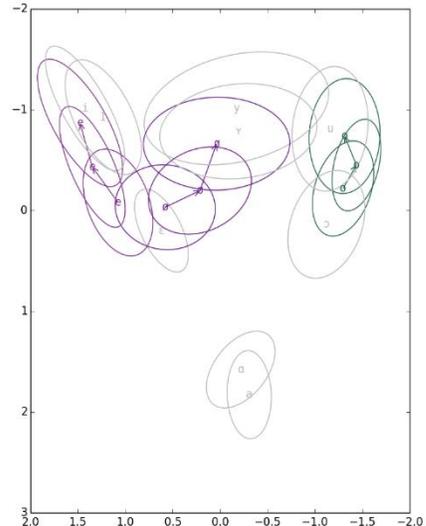


Figure 3.7 *F1 (y-axis) and F2 (x-axis) (normalized) values for three Dutch long mid vowels, as realized by Spanish learners (25%, 50%, 75%); the position of the nine Dutch monophthongs produced by Spanish learners is indicated in gray; the mean values are indicated by the vowel symbol.*

The vowel ellipses displayed in Figure 3.6 show that all three long mid vowels (/e/, /ø/, /ɛ/) are slightly diphthongized in native Dutch. The trajectories of the long mid vowels are towards the three high vowels (/i/, /y/, /u/). MANOVAs were conducted on the three pairwise combinations of the three

Dutch long mid vowels, as produced by native Dutch speakers. All native Dutch vowel realizations are significantly different in their F1 and F2 values at 50%. Using Pillai's trace results in values of above .900 for 50% and for all three time points (25% 50%, 75%) as predictors.

Figure 3.7 displays the ellipses representing the realizations of the Dutch long mid vowels produced by Spanish learners. The ellipses of the Spanish learners are rather spacious, larger than those of native Dutch speakers, but with a clear spectral time trajectory. It shows that Spanish learners diphthongize the three Dutch long mid vowels, and that their vowel realizations resemble the Spanish vowel combinations as earlier advanced in prediction 3. Spanish learners appear to have problems concerning the proper use of the F2 dimension. For example, the learner realizations of /e:/ and the initial part of the trajectory of /ø:/ are more fronted, whereas those of /o:/ and the end point of the trajectory of /ø:/ are more back. The front rounded vowel /ø:/ has a striking trajectory, starting at a high F2 value (left on the *x*-axis, 25%) and ending in a rather low F2 value (right on the *x*-axis, 75%) towards /*[y≈y]*/. The ending of the trajectory of the front rounded /ø:/ towards /*[y≈y]*/ seems to indicate, in line with prediction 2, that a new front rounded category has been created.

To test the vowel distinctions (at 50%), we computed Pillai's trace values for the three vowel pairs. The pair /e:/-/ø:/ clearly has the lowest value for Pillai's trace (.697). The two other pairs, /ø:/-/o:/ and /e:/-/o:/, have values of above .800. Using three time points (25%, 50%, 75%) does not alter the values substantially. A slight increase is found in the pair /e:/-/ø:/ (F1 and F2 only .697, increasing to .703 when including three time points). These relatively high outcomes show that Spanish learners succeed in making distinctions among the three Dutch long mid vowels, although their realizations do not match the native realizations.

3.4.2.3 Diphthongs

The ellipses reflecting the realizations of the Dutch diphthongs (/ei, œy, ou/) by native Dutch speakers and Spanish learners are presented in Figure 3.8 and Figure 3.9 respectively. Three ellipses representing three points of measurements (25%, 50%, 75%) are displayed for each diphthong.

The vowel ellipses displayed in Figure 3.8 show the spectral trajectories of all three diphthongs in native Dutch. Their trajectories move towards the three high vowels (/i, y, u/) and proceed from higher F1 starting points (25%) towards much lower F1 target values (75%). Figure 3.8 shows that the ellipses reflecting the initial part of every diphthong start in the lower area of the vowel space.

MANOVAs were conducted on the three diphthong pairs. All native diphthongal realizations are significantly different in their F1 and F2 values at 50%. Using Pillai's trace results in values of above .777 (clearly the lowest value found for /ei-/œy/) when having F1 and F2 at 50% and above .894 with all three time points (25%, 50%, 75%) as predictors. Using three time points appears to increase the proportion of explained variance in all diphthongal pairs, /ei-/œy/ again having the lowest value.

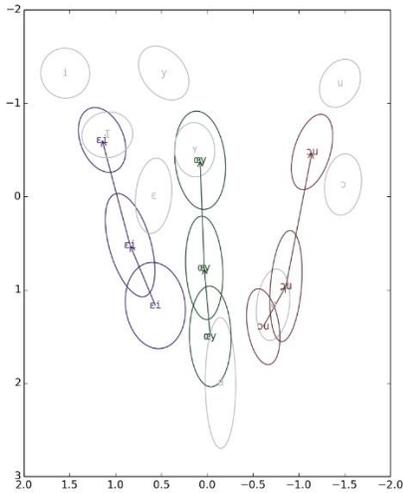


Figure 3.8 *F1 (y-axis) and F2 (x-axis) (normalized) values for three Dutch diphthongs (25%, 50%, 75%) and the nine Dutch monophthongs (in gray) as produced by native Dutch speakers; the mean values are indicated by the vowel symbol; the values for Dutch vowels were drawn from Van der Harst (2011).*

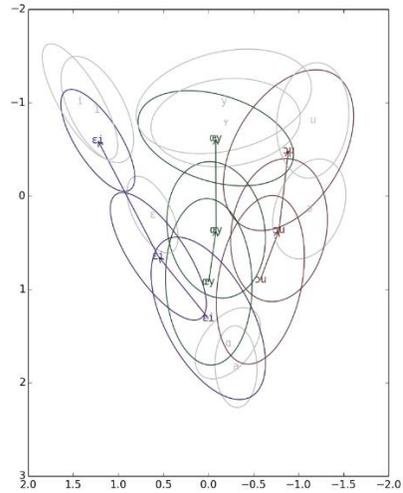


Figure 3.9 *F1 (y-axis) and F2 (x-axis) (normalized) values for three Dutch diphthongs (25%, 50%, 75%) and the nine Dutch monophthongs (in gray) as produced by Spanish learners; the mean values are indicated by the vowel symbol.*

Figure 3 clearly shows that the ellipses reflecting the vowel spaces of the realizations by Spanish learners are larger than those by native Dutch speakers (see Figure 3.8). The production pattern in the learner realizations of Dutch diphthongs would seem to suggest that they employ a wider range to realize the diphthongs. Difficulties in using the F2 dimension properly recur in the diphthongs. Problems in realizing the central front rounded vowel /œy/ are visible from the spacious ellipses reflecting their formant trajectories,

especially at the end point (towards $[\gamma \approx y]$); see Figure 3.9). The trajectory of the $/\text{œy}/$ moving towards the new vowel $[\gamma \approx y]$ appears to show, in agreement with prediction 2, that a new category has been created. The productions of $/\text{ei}/$, as realized by Spanish learners, seem to be more central, whereas those of $/\text{œy}/$ and $/\text{ɔu}/$ occupy a large central-back area of the vowel space. The $/\text{œy}/$ - $/\text{ɔu}/$ confusion is evident from the strong overlap of their ellipses (see Figure 3.9). The learner productions of $/\text{œy}/$, covering a large portion of the vowel space of $/\text{ɔu}/$, indicate that the front rounded $/\text{œy}/$ tends to be produced by Spanish learners as a back vowel, as previously found in Burgos et al. (2014a). All ellipses measured at 25% of the vowel duration seem to start at the Spanish $/a/$, and continue their trajectories towards the high vowels $/i/$, $[\gamma \approx y]$, $/u/$ respectively. These production patterns strongly agree with those formulated in prediction 3, as Spanish learners seem to use Spanish vowel combinations as equivalents for Dutch diphthongs.

The outcome of our statistical analyses (at 50%) show that the vowel pair $/\text{ɔu}/$ - $/\text{œy}/$ distinctly has the lowest value for Pillai's trace (.362), showing that these vowels are not clearly distinguished by the learners. The pair $/\text{ei}/$ - $/\text{œy}/$ has a value of .578, whereas the highest value corresponds to the $/\text{ei}/$ - $/\text{ɔu}/$ distinction (.850). Using all three time points (25%, 50%, 75%) as dependent variables does not substantially improve the Pillai's trace values for the diphthong pairs. These values indicate that Spanish learners are able to distinguish the Dutch diphthongs, although their productions do not match those of the native Dutch speakers, particularly not in the case of the pair $/\text{ɔu}/$ - $/\text{œy}/$.

3.5 Discussion

This study sought to investigate whether the Dutch vowels produced by adult Spanish learners acoustically match those produced by native Dutch speakers, and whether Spanish learners' realizations are influenced by the properties of Spanish L1 vowels. We formulated five predictions based on previous empirical studies on the Spanish L1-Dutch L2 pair and on Flege's (1995) SLM. Our results corroborate all five predictions formulated in Section 2 of this article.

In line with prediction 1, Spanish learners were shown to use their five Spanish core vowels /i, e, a, o, u/ as equivalents for five Dutch monophthongs /i, ε, a:, ɔ, u/ (see Figure 3.5) (cf. equivalence classification in Flege (1995) and Full Copying hypothesis in Escudero (2005)). Establishing this match leaves four monophthongs to be learned, namely /ɣ, y, ɪ, ʌ/. Spanish learners need to apply subtle spectral distinctions to produce the monophthongs /ɪ, ʌ/. If they fail to do so, they may instead resort to duration to produce these subtle distinctions, as formulated in prediction 4. Our outcomes show that the acoustical properties of the learner realizations of /ɪ, ʌ/ largely overlap with those of /i, a:/ indicating that learners could not make the spectral distinctions required for the /ɪ/-/i/ and /ʌ/-/a:/ contrasts. These findings are in line with Burgos et al. (2013, 2014a, b), and with the hypotheses advanced in Flege's (1995) SLM, showing that L2 phones which are similar to L1 categories (Spanish /i, a/) are not likely to be produced in a native-like manner, especially in the case of L2 phones that are non-contrastive in the L1.

In accordance with prediction 2, the results of our acoustical analyses show that Spanish learners were able to produce front rounded vowels, although this newly established category concerns a single but undifferentiated front rounded vowel (see Figure 3.5), defined by us as /[ɣ≈y]/. In line with Flege's

(1995: 240) SLM, the acquisition of a new vowel category might have been triggered by the fact that the vowels /ɣ/ and /y/ do not have a counterpart in the Spanish vowel inventory (cf. Goudbeek et al., 2008). However, the new category is too large and diffuse, as the vowel space used encompasses at least two Dutch vowels, namely /ɣ/ and /y/. Our data indicate that the vowels /y/ and /ɣ/ are primarily confused with each other, but also, although clearly less, with the Dutch back rounded vowel /u/ (counterpart of the Spanish /u/) (see Figure 3.5 and Table 3.6). These findings corroborate the results of Burgos et al. (2014a) which point to a clear /y/-/ɣ/-/u/ confusion.

Prediction 3 stated that adult Spanish learners will overdiphthongize the Dutch long mid vowels (/e:, ø:, o:/) and diphthongs (/ei, œy, ou/), by using Spanish vowel combinations as equivalents. In consonance with prediction 3, the acoustical measurements convincingly show that Spanish learners strongly diphthongized the Dutch long mid vowels and diphthongs, and that their productions resemble Spanish vowel combinations (cf. equivalence classification in Flege (1995) and Full Copying hypothesis in Escudero (2005)), an outcome that fits the observations in Burgos et al. (2014a). The learner productions show outspoken spectral time trajectories along the F1 dimension, although these do not match those of the native speakers (see Figures 3.6 and 3.7 for the long mid vowels, and Figures 3.8 and 3.9 for the diphthongs). The learner spatial trajectories seem to be longer than those of native Dutch speakers, suggesting that Spanish learners are trying to connect two vowels, just as they are used to doing in Spanish vowel combinations. For example, the learner productions of the target long mid vowels /e:, ø:, o:/ have their starting point at the Spanish /e/ (for /e:, ø:/) and /o/ (for /o:/), and their end point at the Spanish /i/ and /u/. Spanish learners may use the new category /[ɣ≈y]/ as end point for the target /ø:/, in combination with the vowel /u/ as the

second vowel to create a vowel combination (see Figure 3.7). The longer trajectories push the three target diphthongs /*ei*, *œy*, *ɔu*/ to the Spanish /*a*/ as the initial vowel, in combination with the high vowels /*i*/ and /*u*/ as closing vowels, leading to productions similar to Spanish vowel combinations /*ai*/ and /*au*/. The new category /*[y≈y]*/, as closing vowel, may help to establish a diphthong meant to correspond to the target diphthong /*œy*/ (see Figure 3.9).

Prediction 4 posits that adult Spanish learners would resort to duration to expand the set of Dutch monophthongs. The feature of duration might be relatively accessible to Spanish learners because their diphthongs have the property of being long. Our results show, in agreement with prediction 4, that learners predominantly use duration to realize the native contrasts /*ɑ*/-/*a:*/, /*l*/-/*i*/ and /*ʏ*/-/*y*/ as /*a*/-/*a:*/, /*i*/-/*i:*/ and /*[y≈y]*/-/*[y≈y:]*/. This suggests that the contrastive use of duration is relatively easy to learn for Spanish learners (cf. Burgos et al., 2014b for the contrast /*l*/-/*i*/), despite the fact that Spanish does not have contrastive vowel length (Hualde, 2005; McAllister et al., 2002). Previous studies on perception showed that native speakers of Spanish use duration to a greater extent than native speakers of English (Escudero & Boersma, 2004) and Dutch (Escudero et al., 2009) in distinguishing perceptually difficult contrasts. The results of our investigation into production are in agreement with these perceptual studies and, moreover, go some way to confirming Bohn's (1995) Desensitization Hypothesis which states that "whenever spectral differences are insufficient to differentiate vowel contrasts because previous linguistic experience did not sensitize listeners to these spectral differences, duration differences will be used to differentiate the non-native vowel contrast" (pp. 294–295). Figure 3.3 convincingly demonstrates that Spanish learners succeeded in implementing the short/long contrast, although durations overall are systematically longer

than those of native Dutch speakers. However, despite their longer durations and the fact that most learner realizations are produced with a duration deviance between 25-100 ms (cf. Nooteboom & Doodeman, 1980), none of the durations puts short vowels in the category of long vowels. The only clear exception is the short vowel /y/, which is systematically realized as long.

Finally, our findings indicate that orthography also has an impact on Dutch L2 vowel production (cf. Escudero & Wanrooij, 2010; Escudero et al., 2014; Escudero, 2015), which concurs with prediction 5. The influence of orthography is particularly noticeable in the lengthening of the Dutch vowels /y, a:, e:, o:/, represented by the digraphs <uu>, <aa>, <ee>, <oo> respectively (see Table 3.4). Spanish learners may have associated the digraphs <uu>, <aa>, <ee>, <oo> with sequences of two Spanish phonemes, i.e., /u+/u/, /a+/a/, /e+/e/, /o+/o/ respectively (Escudero & Wanrooij, 2010: 349). Similarly, it is conceivable that the lengthened realizations of the short vowels /i/ and /u/ might have been triggered by the digraphs <ie> and <oe> respectively, as Spanish learners might have interpreted that a digraph representing a single Dutch phone indicates that the phone in question should be lengthened.

We analyzed the Spanish learner data as consisting of one cohesive system, by pooling the data of all 28 participants involved using read speech. Read speech has the disadvantage of being susceptible to orthographic influences, but the support for our predictions 1 to 4 point to a primary impact of phonological features and the strength of L1 entrenchment, orthography being relevant nevertheless, but at a secondary level. The Dutch proficiency levels of our learners varied, as explained in Section 3, as well as their level of proficiency in other languages, but they all had substantial problems with Dutch pronunciation (see Table 3.3 for the correct scores obtained in Burgos et al. (2015)). Pooling the vowel pronunciations of a group of learners allowed

us to generate an overview of pronunciation variants and their distribution in the vowel space.

The results discussed here are in line with cross-language speech perception models which describe the interference of L1-specific sound patterns on the acquisition of L2 speech sounds and specify that L2 sound contrasts that are mapped onto a single native category are the most difficult to learn (Best, 1995; Escudero, 2005; Flege, 1995). As advanced in Flege's (1995) SLM, Spanish learners were able to establish new vowel categories, although these differed substantially from native vowel categories.

3.6 Conclusions

The present study investigated Dutch vowel production accuracy (for all 15 Dutch full vowels) by adult Spanish learners on the basis of acoustical measurements of duration and spectral features. Our results give a clear answer to the general research question: *Do the Dutch vowels produced by adult Spanish learners acoustically match those of native Dutch speakers?* The Dutch vowels produced by adult Spanish learners do not match, in terms of either duration or spectral values, those produced by native Dutch speakers, due to pervasive L1 constraints. The average durations of the learner realizations are consistently longer than those of the native realizations. Learners nevertheless make a distinction between short and long vowels that resembles native Dutch distinction, with the exception of the short vowel /y/, which is consistently produced with a duration typical of long vowels. With respect to vowel spectral values, adult Spanish learners fail to produce the subtle spectral differences required to distinguish Dutch vowel contrasts based on vowel height (e.g., /I/-/i/ and /ʏ/-/y/). As a result, learners predominantly resort to duration to realize these contrasts in a non-native manner, as in the case of the native contrasts /ɑ/-/a:/, /I/-/i/ and /ʏ/-/y/. Our outcomes have

shown that the Dutch vowels which least match the native realizations are the nine monophthongs, in particular vowel contrasts based on spectral features (e.g., /ɔ/-/u/, /ʏ/-/y/, /ɛ/-/ɪ/, /ɪ/-/i/ and /ɑ/-/a:/). Conversely, adult Spanish learners do succeed in employing diphthongization when realizing the Dutch long mid vowels and diphthongs, although their vowel productions exhibit a clear Spanish-like diphthongization (i.e., combining two full vowels).

A particularly relevant outcome of our investigation is that Spanish learners were able to establish a new single L2 vowel category (/ [ʏ≈y] /) encompassing two L2 front rounded vowels (i.e., /ʏ/ and /y/), whereas such vowels do not exist in their native phonology. Overall, our findings suggest the interaction of two mechanisms in adult L2 vowel acquisition processes: the formation of new vowel categories and equivalence classification, as proposed by Flege (1995).

4 *Auris Populi*: Crowdsourced native transcriptions of Dutch vowels spoken by adult Spanish learners

This chapter has been reformatted and slightly modified from:

Burgos, P., Sanders, E., Cucchiarini, C., Van Hout, R., & Strik, H. (2015). *Auris Populi*: Crowdsourced native transcriptions of Dutch vowels spoken by adult Spanish learners. *Proceedings of Interspeech 2015*, Dresden, Germany, 2819–2823.

4.1 Introduction

Studies on second language (L2) acquisition have shown that adult learners seldom achieve a native-like pronunciation (Birdsong & Molis, 2011; Long 1990). Accented speech does not necessarily impede communication as long as the pronunciation of the L2 learners is intelligible and native listeners are able to understand the intended message (Derwing & Munro, 2005). How can we determine whether accented speech is intelligible? Many studies relied on evaluations of experts. Another approach is to use non-expert native listeners to judge non-native speech, sometimes even asking them to evaluate specific phonetic contrasts. These approaches, however relevant, cannot answer the question what native listeners hear and perceive when they listen to accented speech. What brings the crowd's ear, the *auris populi*, when that ear has to listen to accented pronunciations of a series of separate words, spoken by a group of L2 learners?

A self-evident manner of finding out whether a word produced by L2 learners has been perceived or understood is by asking native listeners to orthographically transcribe the words uttered by L2 learners. A strong reason for doing this is that learners do not actually communicate with a limited

number of experts, but with a diverse and extensive group of native listeners. A promising way of reaching this group is by crowdsourcing. In doing so, we will not only obtain a large and diverse group of native listeners, but at the same time we will be able to collect a variety of transcriptions on the speech of many L2 speakers (Eskenazi, Levow, Meng, Parent, & Suendermann, 2013; Parent & Eskenazi, 2011).

The aim of the current study is to investigate how the *auris populi*, the crowd's ear, would deal with possibly deviant L2 vowel realizations. The listeners' judgments revealing the “wisdom of the crowd's ear” (Kunath & Weinberger, 2010) will help us understand which features of the learner vowel productions may cause confusions in non-expert Dutch listeners' perception.

In the remainder of this paper, we first present the research background in Section 2. Section 3 describes the method, the crowdsourcing experiment and the quality control. The results are presented in Section 4 and discussed in Section 5. Finally, we draw the conclusions of our study in Section 6.

4.2 Research background

There are considerable differences between the Dutch and the Spanish vowel inventories (Burgos, Cucchiarini, Van Hout, & Strik, 2013, 2014a; Burgos, Jani, Cucchiarini, Van Hout, & Strik, 2014b, submitted a; Escudero, Benders, & Lipski, 2009; Escudero & Wanrooij, 2010). First, Spanish has five vowels (/a, e, i, o, u/) (Hualde, 2005), whereas Dutch has 15 unreduced vowels (five lax vowels: /ɪ, ɛ, ɔ, ʏ, ʌ/; seven tense vowels: /i, y, u, e:, ø:, o:, a:/; three diphthongs: /ɛi, œy, ɔu/) and the reduced vowel schwa /ə/ (Booij, 1995). Second, Dutch has a lax/tense distinction, including vowel length (short vowels: /i, y, u, ɪ, ʏ, ɔ, ɛ, ʌ/, long vowels: /a:, e:, ø:, o:, ɛi, œy, ɔu/), whereas Spanish does not have contrastive vowel length. Third, Dutch has four front

rounded vowels: /y, y, ø:, œy/, whereas in Spanish all rounded vowels (/o, u/) are back.

Previous research has investigated the speech production of adult Spanish learners of Dutch (Burgos et al., 2013, 2014a, b, submitted a). Studies conducted by Burgos et al. (2013, 2014a) based on samples of extemporaneous speech showed that vowel errors were more frequent and persistent than consonant mispronunciations. For this reason, follow-up research was conducted on the vowels. Burgos et al. (2014b, submitted a) reported on studies in which elicited material containing read speech was employed. The use of read speech containing all speech sounds that are problematic for Spanish learners, was aimed at obtaining sufficient mispronunciations to be acoustically analyzed. Burgos et al. (2014b) studied the production of three vowel contrasts (/ɑ/-/a:/, /ɪ/-/i/, /ɣ/- /ø:/), and that of the Spanish learners' realizations of all 15 Dutch vowels (Burgos et al., submitted a). Both studies (Burgos et al., 2014b, submitted a) concentrated on the acoustic analysis of the vowels produced by the Spanish learners in comparison to those produced by native Dutch speakers, and concluded that adult Spanish learners do not employ duration and spectral properties in a native-like manner. Moreover, in Burgos et al. (2013, 2014a, b, submitted a) it was found that the L1 phonology influences L2 vowel production and that the five Spanish vowels appear to function as “attractors” for the larger set of Dutch vowels. Based on the results of the studies mentioned above, we can advance the following predictions. First, we hypothesize that non-expert native Dutch listeners will transcribe the tokens produced by the Spanish L2 learners differently from their canonical forms. Second, we expect to find the “attractor” effect phenomenon in the listeners' transcriptions. Third, we predict that deviant patterns found in the acoustic measurements on the same speech material will be mirrored in the listener's transcriptions.

4.3 Method

4.3.1 *Speakers*

To obtain a representative sample of Spanish L1-Dutch L2 vowel pronunciation errors, speech samples from 28 adult Spanish learners of Dutch (9 males, 19 females) with varying degrees of proficiency (A1, n=10; A2, n=7; B1, n=4; B2, n=7, according to the Common European Framework of Reference for Languages (CEFR); Council of Europe) were used in the current study. These data had previously been analyzed in Burgos et al. (submitted a).

4.3.2 *Speech stimuli*

The speech stimuli consisted of separate words in Dutch read by adult Spanish learners. Every speaker read a set of 29 monosyllabic words in which all 15 Dutch vowels in stressed position were presented. The same elicitation material was previously used in Van der Harst (2011), and Van der Harst, Van de Velde, and Van Hout (2014). All the words ended either in /s/ or /t/, as it is known that these consonants scarcely alter the quality of the preceding vowel (Van der Harst, 2011; Van der Harst et al., 2014).

Table 4.1 *Selected -s and -t words used as speech stimuli from Van der Harst (2011); Phon = phonological representation (in IPA), Orth = orthographic representation.*

Vowel	Vow	Orth	s-word	t-word
Monophthongs	/i/	<ie>	<i>kies</i> /kis/	<i>riet</i> /rit/
	/ɪ/	<i>	<i>vis</i> /vɪs/	<i>fit</i> /fɪt/
	/ɛ/	<e>	<i>zes</i> /zɛs/	<i>vet</i> /vɛt/
	/y/	<uu>	-	<i>fuut</i> /fyt/
	/ʏ/	<u>	<i>zus</i> /zʏs/	<i>put</i> /pʏt/
	/u/	<oe>	<i>poes</i> /pus/	<i>voet</i> /vut/
	/ɔ/	<o>	<i>vos</i> /vɔs/	<i>vlot</i> /vlɔt/
	/ɑ/	<a>	<i>gas</i> /xas/	<i>rat</i> /rat/
Long mid vowels	/aː/	<aa>	<i>aas</i> /aːs/	<i>staat</i> /staːt/
	/eː/	<ee>	<i>mees</i> /meːs/	<i>beet</i> /beːt/
	/øː/	<eu>	<i>neus</i> /nøːs/	<i>neut</i> /nøːt/
Diphthongs	/oː/	<oo>	<i>boos</i> /boːs/	<i>boot</i> /boːt/
	/ɛi/	<ij>	<i>ijs</i> /ɛis/	<i>spijt</i> /spɛit/
	/œy/	<ui>	<i>huis</i> /hœys/	<i>fluit</i> /flœyt/
	/ou/	<ou>	<i>kous</i> /kous/	<i>fout</i> /fɔut/

Table 4.1 shows an overview of all 15 Dutch vowels and their corresponding orthographic and phonological representation. No example of the vowel /y/ followed by /s/ was included, as this combination does not appear in Dutch monosyllabic words, except proper names.

For this experiment we used a set of 29 words produced by 28 Spanish learners. Six speech samples were left out. During the task transcribers were offered a word they had transcribed earlier every 30th token. This was done to calculate the intra-transcriber agreement. The inclusion of repeated items gave a maximum of 833 speech stimuli used in the transcription task.

4.3.3 Listeners

Prior to participating in the experiment, listeners read the instruction of the transcription task. They were told that they were going to listen to utterances and that they literally had to transcribe what they heard using orthographic spelling. Listeners were allowed to transcribe foreign and non-existing words which might closely represent the heard utterance. An online questionnaire was administered to obtain background information about the listeners. The number of questions presented in the questionnaire was limited to keep the crowdsourcing experiment as simple and accessible to non-expert listeners as possible. The online questionnaire contained questions concerning mother tongue, gender, age and completed education. Almost 200 listeners participated in the transcription task. Part of the participants were filtered out, resulting in 159 listeners whose data was included in the current study (see Section 4.3.5). All participants were non-expert native Dutch listeners.

4.3.4 The crowdsourcing experiment

A web application was developed in Django, in which participants could listen to the stimuli and type what they heard. The application was set up in such a way that it was easy to use and also fun to do. Each participant received a score indicating the percentage of “correct” transcriptions. This score was based on the most frequent transcriptions given to a word by all (previous) transcribers. The idea behind providing a score was to motivate the participants and introduce a game element, as the score could be shared on Facebook. This helped recruiting new participants. Participants transcribed 100 tokens on average (see Sanders, Burgos, Cucchiari, & Van Hout (2016) for a detailed description of the application).

4.3.5 *Quality control*

Several criteria were used to filter the data. Only listeners who had Dutch as a native language were included. Secondly, listeners had to transcribe >10 tokens, to be sure that they really got started to perform the task. The maximum of 833 transcriptions per listener was included (three listeners continued to perform a second round).

We used two additional quality control criteria to ascertain the reliability of the data, a measure of intra-transcriber agreement and a measure of inter-transcriber agreement (Eskenazi et al., 2013; Parent & Eskenazi, 2011). The intra-transcriber agreement was based on the transcriptions of the repeated items. The inter-transcriber agreement criterion was based on the percentage of shared common transcriptions (cf. Sanders et al., 2016). Listeners failing to meet both agreement criteria were removed from the database. Filtering our data resulted in a total of 17.534 tokens transcribed and 159 listeners.

4.4 Results

4.4.1 *Listeners' transcriptions, vowel confusions*

The listeners' transcriptions show that both consonants and vowels were given canonical and non-canonical transcriptions. We will now focus on the vowels, although consonants also deserve further investigation. Table 4.2 displays the most frequent listeners' transcriptions per vowel. The 15 target Dutch vowels are presented in alphabetic order in the columns, except for the last three vowels, corresponding to the three diphthongs. The rows show the transcribed vowels, including both the canonical transcriptions of the target vowels (indicated by the black squares) and the non-canonical transcription <ai>. The percentages in the cells indicate how often a transcription was given to a target vowel. The column Total shows the sum of all percentages of transcribed vowels per row. Transcriptions containing percentages of less than 1% are

aggregated in the Rest category (see last row in Table 4.2). Overall percentages for canonical and non-canonical transcriptions were calculated. Our results indicate that 67.44% of all transcriptions are canonical and 32.56% non-canonical.

The various non-canonical transcriptions (see rows in Table 4.2) show that there is variation in the way the vowels were transcribed by the non-expert native Dutch listeners. The highest variation was found in the long mid vowel <eu> and the diphthong <ui>. The lowest variation appears in the vowel <aa>.

An interesting confusion pattern is found in the non-canonical transcriptions for the target vowels <u> and <uu>, which are often confused with each other, and especially with the vowel <oe>, as displayed in Table 4.2.

Table 4.2 shows that the target long mid vowel <ee>, and the target diphthongs <ij> and <ui> have non-canonical transcriptions, such as <ei>, <ai> and <au>, respectively. These transcriptions seem to point to strong diphthongization, as observed earlier in Burgos et al. (2014b, submitted a).

The column Total in Table 4.2 shows that some vowels were more often transcribed by the listeners, namely, <aa>, <e>, <ie>, <o>, <oo> and <oe>, all of them producing percentages above 100. These vowels seem to resemble the five Spanish vowels <a>, <e>, <i>, <o>, <u>, suggesting the idea of the Spanish vowels functioning as “attractors” for the larger set of Dutch vowels, as previously observed in Burgos et al. (2013, 2014a, b, submitted a). A conspicuous case, which needs to be further examined, is the one of the two Dutch vowels <o> and <oo>, which appear to be attracted both by the Spanish vowel <o>.

In order to better understand how non-expert native Dutch listeners cope with transcribing specific vowels in a contrast, we decided to study three Dutch vowel pairs <a>-<aa>, <i>-<ie> and <u>-<eu> in more detail. These

vowels, produced by Spanish L2 learners, were acoustically analyzed in Burgos et al. (2014b). They differ from each other in the way duration and place of articulation are used to make a contrast. The contrast <a>-<aa> is based on duration and place. The distinction between the vowels in the pair <i>-<ie> hinges on place and not on duration, as both vowels are short in native Dutch. The contrast <u>-<eu> is only based on duration, as both vowels have a similar place of articulation and are both front rounded vowels.

Table 4.2 *Most frequent orthographic representations of all 15 Dutch vowels transcribed by non-expert native Dutch listeners; transcribed vowels <1% are aggregated in the Rest category, >10% in grey, >5% in light grey, canonical transcriptions in black squares, the orthographic representation of the target Dutch vowels in the columns, the transcribed vowels in the rows; Vow=Vowel.*

Vow	a	aa	e	ee	eu	i	ie	o	oo	oe	u	uu	ij	ou	ui	Total
a	72.21	13.80	0.25	0.00	0.08	0.00	0.17	2.22	0.74	0.08	1.17	0.00	0.08	1.71	0.08	92.59
aa	23.04	79.12	0.00	0.00	0.08	0.00	0.17	0.00	1.24	0.00	0.00	0.00	0.08	0.81	0.17	104.71
e	1.58	0.67	89.38	18.71	2.18	3.28	1.51	0.08	0.08	0.17	2.33	0.00	4.23	0.32	0.67	125.19
ee	0.00	0.00	1.40	59.31	1.01	1.34	6.54	0.00	0.00	0.00	0.08	0.17	4.39	0.00	0.00	74.24
eu	0.00	0.00	0.08	0.25	65.86	0.08	0.17	0.00	0.00	0.08	0.42	4.78	0.24	3.57	3.76	79.29
i	0.00	0.00	4.94	2.09	0.42	51.34	12.66	0.16	0.00	0.00	2.25	0.17	0.08	0.08	0.17	72.18
ie	0.00	0.00	0.25	2.42	0.67	37.98	73.09	0.00	0.00	0.00	0.42	0.34	0.00	0.00	0.00	115.17
o	0.08	0.00	0.00	0.07	0.25	0.08	0.00	86.34	16.29	2.27	1.33	0.34	0.08	0.49	0.25	107.87
oo	0.00	0.00	0.00	0.00	5.95	0.00	0.08	4.94	72.70	11.09	1.50	5.29	0.00	15.84	1.25	118.64
oe	0.08	0.00	0.00	0.08	0.25	0.34	0.00	2.39	4.38	78.24	34.89	40.96	0.00	3.33	1.75	166.69
u	0.50	0.17	1.23	0.00	2.85	3.36	0.17	0.91	0.00	1.93	44.96	4.61	0.00	1.22	0.50	62.41
uu	0.00	0.00	0.00	0.00	0.50	1.26	0.17	0.00	0.33	1.66	9.16	41.10	0.08	0.08	1.09	55.43
au	0.08	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.08	0.00	0.00	0.00	0.00	1.95	5.10	7.37
ai	0.08	0.42	0.00	0.08	0.59	0.00	0.25	0.00	0.00	0.00	0.00	0.00	9.36	0.00	0.08	10.86
ei	0.00	0.00	1.07	6.68	0.25	0.08	1.17	0.00	0.00	0.00	0.00	0.00	3.17	0.00	0.17	12.59
ij	0.00	0.42	0.33	6.35	0.34	0.00	0.42	0.00	0.00	0.21	0.00	0.00	66.48	0.00	0.50	75.05
ou	0.17	0.51	0.00	0.00	0.84	0.00	0.00	0.33	0.58	2.18	0.08	0.17	0.08	59.46	3.59	67.99
ui	0.92	0.42	0.08	0.17	3.27	0.00	0.92	0.74	0.08	0.17	0.17	0.51	3.66	7.23	72.10	90.44
Rest	1.26	4.47	0.99	3.79	14.61	0.86	2.51	1.73	3.50	1.92	1.24	1.56	7.99	3.91	8.77	59.11

4.4.2 Listeners' transcriptions of three vowel pairs

The listener's transcriptions of the vowels in the pairs <a>-<aa>, <i>-<ie> and <u>-<eu> present different patterns (see Table 4.2). Regarding the pair <a>-<aa>, it appears that the target vowel <a> was more often transcribed as <aa> than the target vowel <aa> as <a>, pointing to an asymmetrical confusion. A similar asymmetrical confusion is found in the transcriptions of the target vowels in the <i>-<ie> contrast. Table 4.2 shows that the target vowel <i> is frequently transcribed as <ie>, more often than <ie> as <i>. Concerning the pair <u>-<eu>, the <u> seems to be overwhelmingly transcribed as <oe> by the listeners, although it also shows other minimal confusions, as displayed in Table 4.2. On the other hand, the target vowel <eu> does not have a clear competitor, and has the highest variation of transcribed vowels of all target vowels (see rows in Table 4.2).

4.5 Discussion

In general, the results of the listener's transcriptions for the pairs <a>-<aa>, <i>-<ie> and <u>-<eu> are in line with the outcomes of the acoustic measurements of the same speech material presented in Burgos et al. (2014b). Asymmetrical confusions between the vowels in the pairs <a>-<aa> and <i>-<ie> have been found in both the listener's transcriptions and in the acoustic data. These are probably due to the existence of Spanish counterparts for the Dutch vowels <a> and <i>. The front rounded vowels in the pair <u>-<eu>, which “are unfamiliar and fall in an empty portion of the native vowel space” (Goudbeek, Cutler, & Smits, 2008, p. 123), exhibit a high degree of variation which can be ascribed to difficulties with front rounding, a phenomenon which does not occur in Spanish (Burgos et al., 2014b). An interesting discrepancy between the listeners' transcriptions and the acoustic data was found with respect to duration. The transcriptions do not indicate longer durations of the

vowels in question, while objective measurements showed that the learners' vowels were longer than the corresponding ones produced by native speakers (Burgos et al., 2014b). This may suggest that native listeners “somehow” normalize duration in learner speech with little consequences for word recognition and intelligibility. However, it is possible that deviant duration values do have consequences for the degree of foreign accent (Derwing & Munro, 2005) that is noticeable in the speech of Spanish learners of Dutch. This is definitely a topic that deserves attention in future research.

The listeners' transcriptions show that some vowels were more often transcribed by the non-expert native listeners, as in the case of <aa>, <e>, <ie>, <o>, <oo> and <oe>. These Dutch vowels resemble the five Spanish vowels <a>, <e>, <i>, <o> and <u>, which confirms the idea of the Spanish vowels functioning as “attractors” for some of the Dutch vowels, as advanced in Burgos et al. (2013, 2014a, b, submitted a). In these studies we already discussed the influence of the L1 orthography on L2 vowel production, which appears to play an important role in the vowel mispronunciations and on the activation of such an “attractor” mechanism.

The influence of the native language is also noticeable in the listeners' transcriptions of the long mid vowel <e> and the Dutch diphthongs <ui> and <eu>. Non-canonical transcriptions of the target vowel <ee> as <ei>, which bears a likeness with the Spanish diphthong <ei>, “exhibiting both a similar degree of formant movement” (Escudero & Williams, 2011: 3), point to diphthongization. Previous studies (Burgos et al., submitted a) already indicated that Spanish L2 learners tend to diphthongize long vowels more than native Dutch speakers. The target vowel <ij> was often transcribed by the non-expert listeners as <ai>. The vowel combination <ai> does not correspond to any vocalic phoneme in native Dutch, but does exist in Spanish and corresponds to the diphthong <ai>. A similar situation applies to the target

vowel <ui>, often transcribed as <au>, which is also a diphthong in Spanish. These findings seem to indicate that the “attractor” effect of the Spanish phonology also involves diphthongs, which are combinations of two vowels in Spanish (Hualde, 2005). The idea of the Spanish diphthongs <ei>, <ai> and <au> functioning as “attractors” for the Dutch long mid vowel <ee> and the diphthongs <ij> and <ui> did not appear from previous studies and can be considered an additional finding brought up by the *auris populi*, the crowd's ear.

As to the value of collecting speech transcriptions through crowdsourcing, we did notice that some of the listeners tended to transcribe what they thought the token was, i.e., the canonical transcription, instead of literally transcribing what they heard, because in this way they could get a high score to share it on Facebook. Possibly, this bias made some listeners not entirely perform the task as we wanted, namely, literally transcribing the tokens spoken by the Spanish L2 learners. However, using only canonical transcriptions would not return a 100% score, as for some words the common transcription was non-canonical. Correctedness scores never were higher than 90%, making the game of transcription hard enough to prevent a single strategy to be successful.

The existence of possible biases is an issue that certainly deserves further examination when dealing with crowdsourced native transcriptions (cf. Eskenazi et al., 2013; Parent & Eskenazi, 2011), but the transcriptions we got seem to reflect quite accurately the phonetic variation in the stimuli. Despite the potential drawback of the *auris populi* methodology, we found crowdsourcing to be a valuable tool to collect a large amount of L2 speech transcriptions from an extensive and diverse group of native non-expert listeners.

4.6 Conclusions

The aim of the current study was to investigate how the *auris populi*, the crowd's ear, would deal with possibly deviant L2 vowel realizations. The transcriptions delivered by the non-expert native Dutch listeners appear to provide a majority of common transcription plus relevant information on variation and details in the way Spanish L2 learners' pronunciation is perceived by non-expert Dutch listeners. The listeners' transcriptions mirror the vowel problems and the “attractor” effect found in previous studies conducted at our lab and based on both expert annotations (Burgos et al., 2013, 2014a) and acoustic measurements (Burgos et al., 2014b, submitted a). The findings of our study confirm that the native human ear is able to perceive deviations in L2 accented vowel realizations. An additional advantage of crowdsourcing is that considerable amounts of speech material from many L2 speakers can be transcribed or rated when many non-expert listeners are willing to participate. We would like to draw the following conclusions. First, the results of our study indicate that Dutch vowels pronounced by Spanish learners were often transcribed differently from the canonical forms by non-expert native Dutch listeners. Second, this study revealed that not only the five Spanish vowels, but also three Spanish diphthongs function as “attractors” for the larger set of Dutch vowels. Third, the listeners' transcriptions of three vowel pairs are in line with results of acoustic measurements of the same speech material, but point to a possibly different role of duration deviations. Fourth, the *auris populi* methodology has proven to be a practical and valuable tool for future L2 speech research.

5 Native listeners' transcriptions of vowel variation in L2 speech

This chapter has been reformatted and slightly modified from:

Burgos, P., Sanders, E., Van Hout, R., Cucchiarini, C., & Strik, H. Native listeners; transcriptions of L2 speech (submitted b).

5.1 Introduction

Research on second language (L2) acquisition has shown that adult learners seldom achieve a native-like pronunciation (Birdsong & Molis, 2001). Interference from the first language (L1) appears to play a decisive role in limiting the production of L2 phones and contrasts (Flege, Schirru, & MacKay, 2003). As a consequence, L2 pronunciations marked by non-native segmentals and suprasegmentals might be hard to process by native listeners, as accented speech or foreign accent may differ substantially from the oral production patterns with which they are familiarized (Munro, Derwing, & Morton, 2006).

Although it is widely accepted that adult learners retain a foreign accent, even when achieving native proficiency in other aspects of L2 production (Flege, Birdsong, Bialystok, Mack, Sung, & Tsukada, 2006), several sociolinguistic studies have shown that a foreign accent can be disadvantageous for successful interaction and social acceptance (Brennan & Brennan, 1981; Lippi-Green, 1997; Moyer, 2013). However, accented speech does not necessarily impede communication provided that the pronunciation of the L2 adult learners is intelligible (Derwing & Munro, 2005). Munro and Derwing (1995) define intelligibility as “the extent to which a speaker’s message is actually understood by a listener” (p. 76). How can we determine whether accented L2 speech is intelligible? Several methods have been

employed in earlier studies. For example, listeners have been asked to orthographically transcribe what they have heard (Bent & Bradlow, 2003; Derwing & Munro, 1997), to identify the phoneme they have heard (Flege, Bohn, & Jang, 1997; Van Wijngaarden, 2011) or to rate intelligibility on a Likert scale (Fayer & Krasinski, 1987). Various studies have relied on the evaluation of experts (Gass & Varonis, 1984; Munro & Derwing, 1995), whereas others have used non-expert native listeners to judge L2 speech, or even to evaluate specific phonetic contrasts (Jułkowska & Cebrian, 2015; Magen, 1998; Van Wijngaarden, 2001). Some studies have taken a mixed approach, employing both experts and non-expert native listeners (Kennedy & Trofimovich, 2008) to rate the intelligibility of L2 speakers.

The advantage of using experts is that they can apply their linguistic knowledge to analyze L2 pronunciations. However, eventually L2 learners in the real world will likely interact with naïve native speakers/listeners when communicating in the L2 and not with a limited group of native experts. For this reason, it is important to involve the evaluation of naïve native speakers/listeners in research that assessed the intelligibility of L2 learners' speech. Asking a diverse group of naïve native listeners to orthographically transcribe L2 learners' speech (Bent & Bradlow, 2003; Derwing & Munro, 1997) can help to determine which features of learner productions may cause intelligibility problems.

The main aim of the present study is to investigate what a large sample of naïve native listeners perceives when listening to monosyllabic, separate Dutch words (containing all Dutch vowels) produced by L2 learners in a reading task. More precisely, we used transcriptions of the speech of Spanish L1 learners of Dutch (henceforth, Spanish speakers) by a diverse group of non-expert native Dutch listeners to uncover the difficulties native listeners experience when perceiving Spanish-Dutch vowels. We sought to determine

whether speakers of Spanish, with a relatively small and simple vowel inventory of five vowels (Hualde, 2005), achieve a high or low degree of intelligibility when producing L2 vowels from a larger and more complex vowel system such as Dutch, which has 15 vowels (Booij, 1995) (see Iverson & Evans (2007) who showed that size and complexity of the L1 vowel inventory affect L2 vowel learning, which can subsequently affect the intelligibility of L2 vowels). A match between the vowel transcribed by the native listeners and the canonical (target) form of the same vowel would indicate that the intended vowel as realized by the Spanish speakers was intelligible for the native Dutch listeners.

Next, we briefly describe the phonological properties of the Spanish and Dutch vowel systems. As already mentioned, Spanish has a straightforward five vowel system (/a, e, i, o, u/) (Hualde, 2005), whereas Dutch has a complex vowel system containing 15 full vowels (monophthongs: /i, y, u, ɪ, ʏ, ə, ε, ɑ, a:/; long mid vowels: /e:, ø:, o:/; diphthongs: /ei, œy, ou/) (Adank, Van Hout, & Smits, 2004; Booij, 1995). Four features characterize the differences between the two vowel systems. Spanish does not have phonemic vowel length (Hualde, 2005), whereas Dutch has a strict lax/tense distinction (lax vowels: /ɪ, ε, ə, ʏ, ɑ/; tense vowels: /i, y, u, e:, ø:, o:, a:/), which crosses the short/long distinction (short vowels: /i, y, u, ɪ, ʏ, ə, ε, ɑ/; long vowels: /a:, e:, ø:, o:, ei, œy, ou/) (Adank et al., 2004). Spanish does not have the feature of front rounding, as all rounded vowels in Spanish are back vowels (/o, u/) (Hualde, 2005), whereas Dutch has four front rounded vowels (/ʏ, y, ø:, œy/). Spanish does not have diphthongs at the phoneme level, that is, single phonemes defined by their trajectory between two vowel positions; instead it has a rich inventory of 14 vowel combinations (Hualde, 2005). Dutch has diphthongs at the phoneme level, such as /ei, œy, ou/. The Dutch long mid

vowels (/e:, ø:, o:/) are not considered to be full diphthongs, but they are slightly diphthongized (cf. Adank et al., 2004; Van der Harst, Van de Velde, & Van Hout, 2014). Finally, Spanish is distinguished by three height values (high: /i, u/; mid: /e, o/; low: /a/) (Hualde, 2005), whereas Dutch is characterized by four height values (high: /i, y, u/; high mid: /ɪ, ʏ, e:, ø:, o:/; low mid: /ɔ, ɛ, ɛi, œy, ou /; low: /ɑ, a:/) (Booij, 1995).

Crowdsourcing is a potentially attractive tool to recruit participants, in this case native listeners: it allows for a diverse group of native listeners to be reached, and, at the same time, for the creation of a large corpus of transcribed speech (Eskenazi, Levow, Meng, Parent, & Suendermann, 2013). For this reason, we conducted an earlier study using crowdsourced sampling to investigate how native listeners would perceive possibly deviant L2 vowel realizations. We asked non-expert native Dutch listeners to orthographically transcribe speech samples of Spanish L1 learners of Dutch (see Burgos, Sanders, Cucchiarini, Van Hout, & Strik (2015) for the results of the crowdsource study, and Sanders, Burgos, Cucchiarini, & Van Hout (2016) for a description of the web application used). We collected data on the native listeners' orthographic transcriptions for all 15 Dutch vowels. The analyses showed that the Dutch vowels realized by the Spanish speakers were often transcribed differently from their canonical forms. The study also highlighted potential methodological drawbacks, such as the lack of direct control over the selection of transcribers. Also, the inclusion of a game element with a score that could be shared on Facebook (aimed at recruiting more transcribers) might have affected the outcomes, as some listeners tended to transcribe what they thought the token was, i.e., to give the canonical transcription, instead of transcribing what they heard, because this increased their score. We therefore decided to conduct a more controlled study, using snowball sampling, to determine the consistency of the outcomes in sampling diverse groups of non-

expert native listeners transcribing non-native speech. Snowball sampling involves recruiting subjects from the social networks of a starting set of individuals. Furthermore, we included native Dutch speech samples in the stimulus set, which could be used as anchor points by the transcribers. Obtaining outcomes on the percentages of canonical transcriptions and most frequent vowel confusions per target vowel similar to those found in our earlier study would support the usefulness of the concept of the crowd's ear to investigate difficulties native listeners have in perceiving specific L2 speech. This could then help to spotlight problematic areas of Spanish speakers' pronunciation of Dutch that could inform pronunciation teaching.

Next, we specifically address investigations related to the production of Spanish-Dutch vowels by Spanish learners. A number of earlier studies by Burgos, Cucchiarini, Van Hout, and Strik (2013, 2014a) and Burgos, Jani, Cucchiarini, Van Hout, and Strik (2014b, submitted) have found that Spanish speakers have difficulties with the production of Dutch vowels. Four specific problems were observed: 1) problems concerning the /ɑ/-/a:/ contrast (based on vowel height and duration), 2) problems with the /ɪ/-/i/ distinction (based on vowel height), 3) confusion problems between the vowels /y/, /ɤ/ and /u/ (based on height and front rounding), and 4) problems related to extreme diphthongization in the case of the long mid vowels and diphthongs. These difficulties in production were mirrored by the native listeners' transcriptions of Dutch vowels spoken by Spanish speakers reported in Burgos et al. (2015). We may thus formulate specific predictions for these four difficulties in the present study.

Prediction 1: The target vowel /ɑ/ will be perceived as /a:/.

Spanish speakers have problems in making the vowel contrast /ɑ/-/a:/ (Burgos et al., 2013, 2014a, b, submitted), most likely because these L2 phones are non-contrastive in the L1, as both resemble the Spanish /ɑ/ (cf.

Flège, 1995). An asymmetrical confusion between these vowels was found by Burgos et al. (2015) as the target vowel /ɑ/ was more often transcribed as /a:/ than the target vowel /a:/ as /ɑ/. Jani, Cucchiarini, Van Hout, and Strik (2015) analyzed the acoustic confusability of the same set of L2 vowels. Interestingly, this vowel pair showed the same asymmetry, though less strong. Our expectation is that the target vowel /ɑ/ will be less intelligible and, therefore, less often perceived by the listeners than the /a:/. The target vowel /ɑ/ will be perceived as /a:/.

Prediction 2: The target vowel /ɪ/ will be perceived as /i/.

Earlier findings by Burgos et al. (2013, 2014a, b, submitted) provided evidence of the difficulties Spanish speakers have in producing the /ɪ-/i/ distinction, most likely because of the resemblance of these vowels to the Spanish /i/ (cf. Flège, 1995). The listeners' transcriptions in Burgos et al. (2015) exhibited an asymmetrical confusion between these vowels, as the target vowel /ɪ/ was often transcribed as /i/, more often than /i/ as /ɪ/. Jani et al. (2015) found the same asymmetry in the acoustic confusability of the vowels /ɪ/ and /i/, but less strong. Previous findings on the /ɪ-/i/ distinction lead to the prediction that the target vowel /ɪ/ will be less intelligible than the /i/. The intended vowel /ɪ/ will often be perceived as /i/.

Prediction 3: The front rounded vowels /y/ and /ʏ/ will be perceived as the back rounded vowel /u/.

Burgos et al. (2013, 2014a, b, submitted) reported that the front rounded vowels /y/ and /ʏ/, which do not occur in Spanish, were difficult to pronounce for Spanish speakers and were often realized as the back rounded vowel /u/ (counterpart of the Spanish /u/). Only back vowels are rounded in Spanish (/o, u/) (Hualde, 2005), which may explain why Spanish speakers tend to realize

the vowels /y/ and /ʏ/ as /u/. The /y/-/ʏ/-/u/ confusion was confirmed in the acoustic studies of Burgos et al. (submitted) and Jani et al. (2015). This brings us to the prediction that the target vowels /y/ and /ʏ/ will be the least intelligible vowels for the native listeners. The front rounded vowels /y/ and /ʏ/ will be perceived as the back rounded vowel /u/.

Prediction 4: The long mid vowels /e:, ø:, o:/ and diphthongs /ɛi, œy, ɔu/ will be perceived with an extreme diphthongization

Previous findings revealed that Spanish speakers use Spanish vowel combinations when realizing the Dutch long mid vowels and diphthongs (Burgos et al., submitted). This production pattern was confirmed by the listeners' transcriptions in Burgos et al. (2015). The target vowel /ɛi/ was remarkably often given the non-canonical transcription /ai/, a vowel combination that happens to occur in Spanish. This shows that Spanish speakers employ diphthongization but seem to resort to Spanish vowel combinations when realizing the long mid vowels and diphthongs. If this indeed applies, native listeners will perceive numerous diphthongs that deviate substantially from their canonical (target) forms, which can be interpreted as extreme diphthongization or overdiphthongization. The prediction is that a low degree of intelligibility will be found for the target long mid vowels and diphthongs. We predict that listeners will perceive and transcribe the overdiphthongization of the Dutch long mid vowels and diphthongs produced by Spanish speakers, and that this overdiphthongization will lead to specific confusion patterns in the non-native vowels. More precisely, the diphthong /ɛi/ will be transcribed by the listeners as the overdiphthongized vowel combination /ai/, which corresponds to the Spanish vowel combination /ai/.

5.2 Method

5.2.1 Listeners

A group of non-expert native Dutch listeners was recruited using snowball sampling. Each individual was asked to recruit subjects from his/her social network, both family members and friends. The individuals themselves did not take part in the experiment. A total of 25 undergraduate students (7 males and 18 females) from the Radboud University Nijmegen (the Netherlands) were found to be willing to recruit a minimum of five Dutch native listeners each. These native listeners had to be 18 years old or older, native Dutch speakers, unfamiliar with Spanish-accented Dutch, and non-expert listeners, i.e., non-linguistically trained listeners who had not studied either Linguistics or Philology. In total, the students recruited 139 non-expert native Dutch listeners for the listening task (see Subsection 2.3). Seven listeners did not complete the task so their transcription data were not used. This resulted in transcription data from 132 non-expert native Dutch listeners (59 males, 73 females).

5.2.2 Speech stimuli

The speech stimuli employed in the current study are part of an existing corpus of Spanish L1 Dutch L2 (cf. Burgos et al., 2014b), which comprises systematic productions of Dutch speech sounds by 28 adult Spanish L1 learners of Dutch (9 males, 19 females). The subset we used consists of a list of 29 words in Dutch the listeners were presented with one by one on a computer screen. The same elicitation materials were used in Van der Harst (2011), and Van der Harst et al. (2014) to describe the Dutch vowel system. The original word list consists of 278 monosyllabic and disyllabic words representing all Dutch vowels in different contexts (cf. Van der Harst (2011) for a detailed description of the word list in Dutch). The subset of 29 monosyllabic words

contains all 15 Dutch vowels in stressed position, as shown in Table 5.1. The words ended either in /s/ or /t/, as it is known that these consonants scarcely alter the quality of the preceding vowel (Van der Harst, 2011, p. 146; Van der Harst et al., 2014, p. 254). No example of the vowel /y/ followed by /s/ was included, as this combination does not occur in Dutch monosyllabic words, except in proper names.

Table 5.1 Selected *-s* and *-t* words as speech stimuli (Van der Harst, 2011); Phon = phonological representation (in IPA), Orth = orthographic representation.

Vowel	Phon	Orth	s-word	t-word		
Monophthongs	/i/	<ie>	<i>kies</i>	/kis/	<i>riet</i>	/rit/
	/y/	<uu>	-	/y/	<i>fuut</i>	/fyt/
	/u/	<oe>	<i>poes</i>	/pus/	<i>voet</i>	/vut/
	/ʌ/	<ɪ>	<i>vis</i>	/vɪs/	<i>fit</i>	/fit/
	/ɪ/	<u>	<i>zus</i>	/zɪs/	<i>put</i>	/pɪt/
	/ɔ/	<o>	<i>vos</i>	/vɔs/	<i>vlot</i>	/vlɔt/
	/ɛ/	<e>	<i>zes</i>	/zɛs/	<i>vet</i>	/vɛt/
	/ɑ/	<a>	<i>gas</i>	/xɑs/	<i>rat</i>	/rat/
	/a:/	<aa>	<i>aas</i>	/a:s/	<i>staat</i>	/start/
Long mid vowels	/e:/	<ee>	<i>mees</i>	/me:s/	<i>beet</i>	/be:t/
	/ø:/	<eu>	<i>neus</i>	/nø:s/	<i>neut</i>	/nø:t/
	/o:/	<oo>	<i>boos</i>	/bo:s/	<i>boot</i>	/bo:t/
Diphthongs	/ɛi/	<ij>	<i>ijs</i>	/ɛis/	<i>spijt</i>	/speit/
	/œy/	<ui>	<i>huis</i>	/hœys/	<i>fluit</i>	/flœyt/
	/ou/	<ou>	<i>kous</i>	/kous/	<i>fout</i>	/fɔt/

Each word was recorded by each of the 28 speakers only once. Of the resulting 812 recordings (29 words x 28 speakers), six speech samples (from

six different subjects) were left out due to defective recording. Overall then, the non-native stimuli consisted of a total of 806 word tokens.

Most of the speakers were born and raised in Spain ($n = 21$), whereas a smaller number ($n = 7$) were originally from different Latin American countries (Argentina, Dominican Republic, Guatemala, Mexico and Venezuela). All speakers were living in the Netherlands, were familiar with the Common European Framework of Reference for Languages (CEFR; Council of Europe) and rated their own proficiency level of Dutch using the CEFR Self-Assessment Grid (Council of Europe). The Spanish speakers in our study assessed themselves at four proficiency levels according to the CEFR, namely, A1 (i.e., CES (= Cambridge English Scales) 100-119), A2 (i.e., CES 120-139), B1 (i.e., CES 140-159) and B2 (i.e., CES 160-180) (UCLES, 2015) (cf. Burgos et al. (submitted) for a more detailed information about the Spanish speakers).

We also used native speech stimuli from previous investigations (cf. Van der Harst, 2011; Van der Harst et al., 2014), namely the same 29 monosyllabic words recorded by two native speakers of Standard Dutch (one male and one female). This resulted in 58 word tokens (29 words x 2 speakers). The purpose of including these native Dutch data as stimuli in the transcription task was to provide listeners with anchor points. Employing both native and non-native data is a method often used in L2 speech perception studies, particularly in intelligibility tests (cf. Bent & Bradlow, 2003; Piske, MacKay, & Flege, 2001). The inclusion of the native data resulted in a total of 864 speech stimuli (806 non-native and 58 native tokens).

5.2.3 Procedure

A web application was developed, in which participants were asked to listen to speech stimuli and type what they heard. The same web application was used in the crowdsourcing study (cf. Burgos et al., 2015; Sanders et al., 2016). The application met several criteria aimed at facilitating task performance:

- The interface was user-friendly. Participants listened to a word. They could click on the play button to listen to the word again. After typing in what they heard, they pressed enter or clicked on OK to hear the next word. Participants were required to type every word until the transcription task was completed.
- The login procedure was user-friendly: participants logged in via Facebook or with an alternative login (username and password).
- The application remembered which tokens a participant had transcribed so they could log out at any time and continue the task where they left off.
- The online questionnaire used to collect the participants' background information was brief.

Prior to participating in the experiment, listeners read the instructions for the transcription task. They were told that they were going to listen to 100 words, consisting of existing and non-existing (i.e., nonsense) words¹, and that they had to transcribe what they heard using Dutch orthography. Listeners were encouraged to type what they heard literally and not what they thought a speaker might have meant or wanted to say.

¹ Native Dutch listeners were told that some of the words they were going to hear were non-existing (i.e., nonsense) words as some of the Dutch words mispronounced by Spanish speakers were likely to be perceived by the listeners as words (or utterances) which did not exist in their native Dutch lexicon.

To become familiar with the transcription task, listeners were presented with an example of how to carry out the task. They heard a word and then looked at the two different ways in which the word/token they heard could be transcribed. Listeners were not told that the utterances were spoken by Spanish learners of Dutch and by native Dutch speakers.

Although listeners were encouraged to complete the task in one go, they were allowed to log out and to complete the transcription task at a later time by logging in again. They were informed that the task would take approximately 30 minutes. After logging in, listeners were asked to complete an online questionnaire aimed at obtaining information about their mother tongue, gender, age, the foreign languages they spoke and their proficiency level in Dutch, profession or studies, and education.

As mentioned above, the stimulus set consisted of a total of 864 speech tokens (806 non-native and 58 native tokens). For the transcription task, listeners were required to transcribe 100 tokens. Each listener transcribed a different set of 100 tokens. Of each set of 100 tokens, randomly presented, 87 corresponded to non-native data, 10 to native data, and 3 to repeated items, i.e., tokens that had previously been transcribed by that particular listener. The repeated items were used to calculate the intra-transcriber consistency. The native tokens were interspersed among the non-native tokens. The frequency with which the 87 non-native tokens (3 sets x 29 words) were presented was counterbalanced; they were presented in three sets containing the 29 different words each (see Table 5.1). Each set was followed by the repeated item from that set, randomly presented.

5.2.4 Intra-transcriber consistency

A total of 439 word tokens (in total) were transcribed more than once. From those 439 word tokens, 420 were transcribed twice, 18 word tokens were

transcribed three times and one word token was transcribed four times. The intra-transcriber consistency was based on the transcriptions of the repeated items. Examining the transcriptions of the repeated items resulted in 387 cases in which the intra-transcriber agreement was consistent (88.2%), and 52 cases in which it was inconsistent (11.8%). Of those 52 cases of inconsistencies in the intra-transcriber agreement, at least two were due to typos. Inconsistencies did not show specific transcription patterns and were distributed over all transcribed vowels. The results of the transcribed vowels are presented below.

5.3 Results

The results of the current study are presented in three subsections. In the first subsection the non-native confusion matrix will be discussed. The second subsection compares the outcomes of the present study with those of our earlier crowdsourcing study. The third subsection considers the native confusion matrix.

5.3.1 *The non-native confusion matrix*

The columns of the non-native matrix in Table 5.2 present the 15 target vowels, corresponding to the nine monophthongs (<ie>, <uu>, <oe>, <i>, <u>, <o>, <e>, <a>, <aa>), the three long mid vowels (<ee>, <eu>, <oo>), and the three diphthongs <ij>, <ui>, <ou>). The percentages in the cells indicate how often a particular transcription was given to a target vowel. The column Total shows the sum of all percentages of transcribed vowels per row, divided by the number of target vowels in the columns. The rows show the transcribed vowels, including the canonical (indicated in green) and non-canonical transcriptions of the target vowels. Notice that the target vowels <ij> and <ou> can also be orthographically represented as <ei> and <au>, respectively. The transcriptions <ei> and <au> were subsumed under <ij> and

<ou>. Most of the listeners' transcriptions turned out to be in Dutch orthography. The non-standard transcription <ai>, representing an open diphthong, is added as a row, as it was the only variant that obtained a frequency in one of the columns of more than 5%. All other transcription variants are subsumed under the Rest categories. The Rest categories include non-canonical transcriptions that are related to longer duration (L), diphthongization (D), other transcriptions (O) and consonants (C).

Overall percentages for canonical and non-canonical transcriptions were calculated. Our outcomes show that 58.37% of all transcriptions are canonical and 41.63% non-canonical.

The highest percentage of canonical transcriptions is found for the vowel <e> (84.38), and the lowest for <uu> (31.75). Ordering all target vowels from the highest to the lowest percentage of canonical transcriptions resulted in the following list: <e>, <o>, <aa>, <oe>, <ie>, <a>, <oo>, <ui>, <ee>, <eu>, <ou>, <ij>, <i>, <u> and <uu> (cf. prediction 3 for <u> and <uu>). The first five target vowels correspond to the five Spanish core vowels. The column Total in Table 5.2 shows that the same five vowels were more frequently transcribed on average than other monophthongs. These higher averages match the attractor effect or similarity attraction of the five Spanish core vowels suggested in Burgos et al. (2013, 2014a, b, 2015, submitted).

The non-canonical transcriptions assigned to the target vowels indicate that the listeners perceived the vowels spoken by the Spanish speakers in various ways. The highest differentiation in transcribed vowels was found in the front rounded vowels <ui> and <eu> with 15 and 14 different vowel categories, respectively, and with more non-canonical transcriptions in the D row (resp. 9.91% and 13.26%; Table 5.2). This variation in perception might be related to the fact that <ui> and <eu> are front rounded vowels which are known to be difficult for Spanish speakers (Burgos et al., 2013, 2014a, b, 2015). The

lowest number of variants is found in the target vowel <i>, with only seven different vowel categories.

Table 5.2 shows that the vowels in the pair <a>-<aa> were hard to distinguish, but that <a> is more often perceived as <aa> (29.78%) than the other way around (13.15%), as advanced in prediction 1. This asymmetry is also found in the <i>-<ie> pair: <i> is perceived as <ie> (43.15%), and <ie> as <i> (12.73%) (see prediction 2).

The non-canonical transcriptions of the target vowels <u> and <uu> provide evidence that these vowels are frequently confused with each other, but especially with the vowel <oe>, meaning that they were perceived as back and round, confirming prediction 3.

Our outcomes show that the native listeners used non-standard Dutch orthography to transcribe the vowels they heard. For example, <aaa>, <aaaa>, <aah> or <aaah> were frequent vowel combinations, indicating that the vowel the listeners heard was produced with an extremely long duration. The percentages in the L row of the Rest category (see Table 5.2) are clearly lower than those of the D row. The listeners employed numerous and various non-standard transcriptions such as <oeie>, <aaj> or <eeuw> to transcribe the vowels they heard. The target vowel with the highest percentage of transcriptions indicating distinct diphthongization was <eu>, followed by <ui> and <ij>. In line with prediction 4, the transcriptions seem to indicate that listeners perceived the diphthongs of the Spanish speakers as extremely diphthongized. This conclusion is supported by the special status of the non-canonical <ai> variant with a score of 15.78% for the <ij>.

Table 5.2 *Most frequent orthographic representations of all 15 Dutch vowels produced by Spanish speakers; target vowels in the columns (canonical transcriptions in green), transcribed vowels in the rows; transcribed vowels < 5% are included in the Rest category; L=Longer duration, D=Diphthongization, O=Other transcriptions, C=Consonant.*

Vowel	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	Total
ie	63.41	0.75	0.00	43.15	0.61	0.00	0.00	0.00	0.00	1.10	0.61	0.00	0.00	0.12	0.12	7.55
uu	0.37	31.75	1.97	0.98	9.88	0.00	0.00	0.00	0.00	0.12	0.37	0.12	0.00	0.61	0.50	2.11
oe	0.12	28.50	69.79	0.12	35.98	1.58	0.25	0.00	0.00	0.00	0.25	4.56	0.00	1.70	3.72	9.12
i	12.73	1.50	0.12	43.89	3.66	0.00	5.38	0.00	0.00	2.82	1.35	0.00	0.25	0.24	0.12	4.93
u	0.49	3.75	2.22	3.79	38.54	2.91	1.63	0.12	0.12	0.00	2.83	0.12	0.00	0.12	2.73	3.99
o	0.00	0.00	4.07	0.00	2.20	79.88	0.00	0.62	0.00	0.00	0.37	18.74	0.00	0.12	1.36	7.48
e	1.98	0.00	0.12	1.59	1.95	0.00	84.38	1.86	0.73	15.71	3.93	0.00	2.96	0.85	0.25	7.93
a	0.12	0.00	0.37	0.00	1.22	5.45	1.13	62.53	13.15	0.25	0.00	1.85	0.00	0.85	1.86	6.10
aa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	29.78	76.86	0.12	0.12	1.23	0.12	0.36	0.99	7.59
ee	11.5	0.25	0.00	4.03	0.00	0.00	1.38	0.00	0.00	52.76	0.86	0.00	5.06	0.12	0.00	5.23
eu	0.00	11.25	0.12	0.00	1.22	0.24	0.25	0.00	0.12	0.25	51.60	0.00	0.00	3.76	3.47	4.60
oo	0.00	15.25	15.66	0.00	1.10	7.15	0.00	0.00	0.00	0.00	7.49	62.64	0.00	4.00	17.84	8.50
ij	1.85	0.25	0.00	0.00	0.12	0.00	1.38	0.00	0.48	13.86	1.60	0.00	53.15	0.85	0.00	5.05
ai	0.12	0.00	0.00	0.00	0.00	0.00	0.63	0.25	0.73	2.94	1.47	0.00	15.78	0.24	0.00	1.53
ui	0.49	0.50	0.37	0.00	0.24	0.24	0.25	1.24	0.37	0.25	3.56	0.00	2.84	57.77	7.68	5.26
ou	0.00	0.25	1.60	0.00	0.49	0.24	0.00	0.12	0.12	0.00	1.23	1.60	0.20	12.26	53.41	4.90
Rest	L	1.11	2.00	0.12	0.12	0.24	0.24	0.38	0.74	1.70	2.70	0.24	2.60	0.00	0.00	0.81
	D	1.23	1.75	1.84	0.00	0.36	0.72	0.77	1.11	1.22	3.05	13.26	4.78	8.01	9.91	3.33
	O	0.00	0.25	0.00	0.12	0.12	0.12	0.13	0.12	0.00	0.36	0.00	0.00	0.12	0.24	0.10
	C	4.45	2.00	1.60	2.20	2.07	1.21	2.13	1.74	4.38	3.68	8.85	1.73	11.59	5.83	2.60

5.3.2 Comparison crowdsource and present study

This subsection compares the transcriptions of the non-native data of the earlier (crowdsource) with those of the present study; for this comparison, the native speech data was removed from the transcription data set of the present study. Table 5.3 compares the percentages of canonical transcriptions per target vowel and the most frequent vowel confusions and their percentages found in the earlier crowdsource study (cf. Burgos et al., 2015) and the present, more controlled, study.

The overall percentage of canonical transcriptions for the crowdsourcing study is 67.80%, which is higher than that of the present study, at 58.82%. A Pearson correlation revealed that there is a significant and extremely strong correlation between the percentages of canonical transcriptions per target vowel found in the two studies ($r = .968$, p (two-tailed) = .000). Furthermore, the pattern going from lower to higher canonical transcription is the same, meaning that there is a convincing degree of concurrence between the outcomes of the two studies.

Table 5.3 presents the most frequent non-canonical transcription category for all target vowels in the two studies. The non-canonical categories in the Conf-CS and Conf-PS rows are the same. An examination of the highest percentages of non-canonical transcriptions in the Rest category of the <eu> and <ui> again showed similarities. For the crowdsourcing study we found the transcriptions <euw> (0.86) for the target vowel <eu> and <auw> (2.05) for the <ui>, whereas for the present study it was <ew> (2.70) for the <eu>, and <aui> (2.06) for the <ui>. The transcriptions of the native listeners in the two studies convincingly indicate strong agreement on the most frequent vowel confusion category per target vowel.

The main difference between the two studies lies in the lower percentages of canonical transcriptions in the current study. We believe that there are two possible explanations. The first relates to the absence of the game element in the present study. As reported in Burgos et al. (2015, p. 2822), the game element in the crowdsourcing study may have led to an artificial increase in the percentage of canonical transcriptions, as some listeners may not have literally transcribed what they heard, but given canonical transcriptions instead, to provide them a higher score. The second potential explanation is that native speech data were included in the transcription task of the current study.

Interspersing non-native with native speech stimuli may have made the native listeners more alert to the task and may have influenced their performance.

Table 5.3 *Percentages of canonical transcriptions per target vowel and most frequent vowel confusions, resulting from the crowdsource and the present study; Vowel = target vowel, CS = crowdsource study, PS = present study, Δ = difference between the percentages of canonical transcriptions of both studies, Conf- = most frequent vowel confusions per study, %Can = average percentage canonical transcriptions.*

Vowel	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	%Can
CS	73.09	41.10	78.24	51.34	44.96	86.34	89.38	72.21	79.12	59.31	65.86	72.70	69.65	72.10	61.41	67.80
PS	63.41	31.75	69.79	43.89	38.54	79.88	84.38	62.53	76.86	52.76	51.60	62.64	53.15	57.77	53.41	58.82
Δ	9.70	9.30	8.40	7.40	6.50	6.40	5.00	9.70	2.20	6.50	14.30	10.10	16.50	14.30	8.00	8.98
Conf-CS	i	oe	oo	ie	oe	oo	i	aa	a	e	Rest	o	ai	Rest	oo	
	12.66	40.96	11.09	37.98	34.89	4.94	4.94	23.04	13.80	18.71	14.61	16.29	9.36	8.77	15.84	
Conf-PS	i	oe	oo	ie	oe	oo	i	aa	a	e	Rest	o	ai	Rest	oo	
	12.73	28.50	15.70	43.15	35.98	7.15	5.38	29.78	13.15	15.71	14.35	18.74	15.78	15.98	17.84	

5.3.3 The native confusion matrix

Table 5.4 presents the percentages of canonical and non-canonical transcriptions given by the native Dutch listeners to the word tokens spoken by two native Dutch speakers. These speakers were assumed to be anchor points, but the scores assigned to their vowel realizations were lower than expected. Our outcomes show that 77.99% of all transcriptions are canonical, whereas 22.01% are non-canonical. The same data from native Dutch speakers were employed in previous investigations (cf. Jani et al., 2015; Van der Harst, 2011; Van der Harst et al., 2014) and no particular anomalies were reported with respect to the speech of these two native Dutch speakers in comparison to the speech of the other native Dutch speakers in the native database.

A thorough examination of the transcriptions of the native data in the present study made clear that the low percentages found were not related to a

specific target vowel, but rather to specific word tokens produced by one of the two native speakers (for instance, the target word “zes”, “six” was often transcribed as “zus”, “sister”).

Table 5.4 shows that the target vowel with the highest percentage of canonical transcriptions is <aa> (96.34%), whereas the target vowel with the lowest is <e> (50.88%).

Most non-canonical transcriptions seem to be related to the specific non-native confusion patterns, e.g. the <i> being transcribed as <ie> (the feature height). Other confusion patterns, associated with the feature of rounding, seem to be distributed over more vowels. Front unrounded vowels were perceived as front rounded vowels: <e> as <u>, <i> as <uu>, and <ij> as <ui>. The feature of rounding is also involved in perceiving front rounded vowels as back rounded vowels: <uu> as <oe> and in perceiving the <a> as <o>. The latter distinction also involves height. It seems that problematic features in perceiving non-native vowels recur in perceiving native vowels.

Table 5.4. *Most frequent orthographic representations of all 15 Dutch vowels produced by native Dutch speakers; target vowels in the columns (canonical transcriptions in green), transcribed vowels in the rows; transcribed vowels < 5% are included in the Rest category; L=Longer duration, D=Diphthongization, O=Other transcriptions, C=Consonant.*

Vowel	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	Total
ie	88.10	0.00	1.30	16.46	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.59
uu	3.57	64.15	0.00	10.13	15.19	0.00	0.00	0.00	0.00	2.90	1.45	0.00	0.00	0.00	0.00	5.17
oe	1.19	22.64	93.51	2.53	2.53	19.44	0.00	0.00	0.00	1.45	1.45	6.58	0.00	0.00	0.00	10.09
i	2.38	0.00	0.00	55.70	0.00	0.93	1.75	0.00	0.00	4.35	0.00	0.00	0.00	0.00	0.00	4.40
u	0.00	5.66	0.00	6.33	78.48	1.85	42.11	0.00	1.22	2.90	1.45	0.00	0.00	0.00	0.00	8.62
o	0.00	0.00	0.00	0.00	0.00	75.93	0.00	21.33	0.00	0.00	0.00	2.63	0.00	0.00	0.00	8.62
e	0.00	0.00	0.00	0.00	0.00	0.00	50.88	0.00	0.00	5.80	0.00	0.00	1.05	0.00	0.00	2.93
a	0.00	0.00	0.00	0.00	0.00	0.00	0.00	74.67	1.22	0.00	0.00	0.00	0.00	1.04	1.64	5.09
aa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	96.34	0.00	0.00	0.00	0.00	0.00	0.00	6.81
ee	0.00	0.00	0.00	2.53	0.00	0.00	0.00	0.00	0.00	72.46	0.00	0.00	1.05	0.00	0.00	4.57
eu	0.00	1.89	0.00	1.27	1.27	0.00	0.00	0.00	0.00	4.35	91.30	0.00	0.00	0.00	0.00	5.95
oo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.45	84.21	0.00	1.04	9.84	6.21
ij	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	70.53	0.00	0.00	5.78
ai	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.45	0.00	0.00	2.11	0.00	0.00	0.26
ui	0.00	3.77	0.00	1.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	20.00	91.67	0.00	9.48
ou	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.22	0.00	0.00	0.00	0.00	1.04	81.97	4.05
Rest	L	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.45	0.00	0.00	0.00	0.00	0.00	0.09
	D	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.45	0.00	1.32	0.00	0.00	0.00	0.26
	O	0.00	0.00	1.30	2.53	1.27	0.00	0.00	1.33	0.00	0.00	2.63	1.05	0.00	0.00	0.69
	C	3.57	1.89	3.90	1.27	1.27	1.85	5.26	2.67	0.00	1.45	2.90	2.63	4.21	5.21	6.56

5.4 Discussion

This study sought to investigate what the crowd's ear perceives when listening to separate Dutch words (containing all 15 Dutch vowels) produced by Spanish speakers. We examined the transcriptions of a large and diverse group of non-expert native Dutch listeners to determine to what extent Spanish-Dutch vowels are intelligible to native listeners, and to uncover the difficulties they experience in perceiving these L2 vowels, which, by extension, indicates potentially problematic areas of production on the part of Spanish speakers.

The results of our study supported our predictions. Consistent with prediction 1, the listeners' transcriptions showed the low intelligibility of the target vowel /ɑ/ which was extensively transcribed as /a:/, more than /a:/ was transcribed as /ɑ/ (see Table 5.2), pointing to an asymmetrical confusion. This confusion pattern in the vowel contrasts /ɑ/-/a:/ concurs with earlier findings of Burgos et al. (2013, 2014a, b, 2015, submitted).

Confirming prediction 2, our outcomes showed that the intelligibility scores of the target vowel /ɪ/ were low (43.89%). The target vowel /ɪ/ was extensively transcribed as /i/, as evident from the highest percentage of non-canonical transcriptions found in the non-native matrix, namely, 43.15% (see Table 5.2). The vowels /ɪ/ and /i/ are non-contrastive in the L1, as both resemble Spanish /i/. Our findings, in line with earlier studies of the same pair of L2 vowels (Burgos et al., 2013, 2014a, b, 2015, submitted), show an outspoken asymmetry in the /ɪ/-/i/ confusion, as attested by the strong over-representation of the vowel /i/ (counterpart of the Spanish /i/).

Acoustic analyses of the vowels in the pairs /ɑ/-/a:/ (based on vowel height and duration) and /ɪ/-/i/ (based on vowel height) revealed that Spanish speakers rely predominantly on duration to make these vowel distinctions (Burgos et al., 2014b, submitted; Jani et al., 2015). The results of the current study seem to suggest that listeners perceived that Spanish speakers resort to duration to make a distinction between a short and a long vowel in the native contrasts /ɑ/-/a:/ and /ɪ/-/i/, as attested by the percentages of these vowels in the L row of the Rest category (resp. 0.74% and 1.70% for /ɑ/ and /a:/, and resp. 0.12% and 1.11% for /ɪ/ and /i/; see Table 5.2). Making a durational distinction is not helpful in perceiving the distinction between /ɪ/ and /i/, which is only based on vowel height in native Dutch. The precise role of duration is

a topic that deserves further attention in future research on Spanish accented Dutch.

In consonance with prediction 3, our results show that the front rounded vowels /y/ and /ɥ/ were the least intelligible vowels for the native listeners, as evident from their rather low percentages of canonical transcriptions, namely, 31.75% and 38.54%, respectively (see Table 5.2). Front rounding is a non-existent feature in Spanish. Difficulties in producing the front rounded vowels /y/ and /ɥ/ have been found in production studies (Burgos et al., 2013, 2014a, b, submitted; Jani et al., 2015).

Prediction 4 posits that low intelligibility would be found for the target long mid vowels (/e:/, ø:/, o:/) and diphthongs (/ɛi, œy, ɔu/), and that the native listeners would perceive and transcribe overdiphthongization in these vowels produced by Spanish speakers. The low intelligibility observed for the long mid vowels and diphthongs, as evident from their low percentage canonical transcriptions (see Table 5.2), is likely to be related to a too strong diphthongization. Our results show that the long mid vowels indeed exhibit a high percentage of special diphthongization transcriptions, especially in the case of the vowel /ø:/ (13.26%) (see D row of the Rest category in Table 5.2). Similar patterns were found for the diphthongs. As expected, the diphthong /ɛi/, and to a lesser degree the long mid vowel /e:/, was often transcribed by the listeners as the overdiphthongized vowel combination /ai/, which corresponds to the Spanish vowel combination /a+/i/. The listeners' transcriptions of /ɛi/ (and /e:/) as /ai/ concur with production patterns, pointing to extreme diphthongization, reported in Burgos et al. (2015, submitted). The transcriptions seem to suggest that Spanish speakers use their rich inventory of Spanish vowel combinations (cf. Hualde, 2005) to produce Dutch diphthongs.

An important implication of the present study is that it shows that the crowd's ear methodology is a useful approach to determine the intelligibility of L2 speech. The findings of the current investigation are in line with those of the earlier crowdsourcing study (Burgos et al., 2015). The native listeners' transcriptions in both experiments clearly reflect agreement on the rank order of the canonical transcription, with an almost perfect correlation, and on the most frequent vowel confusions per target vowel (see Table 5.3). The consistency of the confusion patterns and the considerable degree of correlation between the outcomes of both studies involving a diverse sample of native listeners corroborate the usefulness of consulting non-expert listeners in L2 speech research on intelligibility.

An unexpected outcome of our study was the low percentage of canonical transcriptions obtained for the native data (77.99%) (see Table 5.4), in combination with the lower percentage of canonical transcriptions in the present study (58.82%) compared to our crowdsourcing study (67.80%). Jani et al. (2015), using the same native database (cf. Van der Harst, 2011; Van der Harst et al., 2014), obtained a percentage of 92.00% correct classifications, a percentage that confirms that the native data did not contain any anomalies. The inclusion of native speech samples in the present study was meant to contribute to the validity of the task performance as native tokens could be used as anchor points by the transcribers. The presence of native samples seems to be a plausible explanation for the lower scores obtained in the current study compared to our crowdsourcing study, although other factors may have contributed too, including the absence of a gaming element in the current study. Another question is how to explain the relatively low scores of the native speech samples. The explanation we would tentatively put forward is that the native listeners' performance was influenced by perceptual adaptation processes because of the input of a large amount of non-native speech (cf.

Bradlow & Bent, 2008; Cutler, 2012). Native listeners may have temporarily adapted or shifted their category boundaries to the ambiguous phonemes in the non-native samples. Such processes could help to understand why native front unrounded vowels were transcribed as front rounded vowels.

Earlier work has shown that native listeners can adjust rapidly to non-native realizations, ignoring their long-term native representations of those realizations (Clarke & Garret, 2004). Our results seem to provide evidence of what Cutler (2012) describes as “the plasticity in adult native listeners’ perception” (p. 375). Native listeners’ perceptual adaptation may generate a boundary shift between categories when being exposed to non-native (cf. Clarke & Luce, 2005) and native speech, especially when vowels share a phonological feature such as rounding (cf. Chládková, Podlipský, & Chionidou, 2017).

The results discussed here are in line with previous research on the intelligibility of L2 speech. Van Wijngaarden (2001) investigated the effect of non-nativeness on the effectiveness of speech communication including native listeners, and concluded that vowels that are difficult for L2 speakers to produce are also difficult for native listeners to recognize. We found that vowels that are difficult for Spanish speakers to realize (cf. Burgos et al., submitted; Jani et al., 2015 for /y/ and /ɣ/) are also difficult for native Dutch listeners to perceive (see e.g., outcomes for /y/ and /ɣ/ in Table 5.2). It is possible that other factors such as the lack of familiarity of native Dutch listeners with specific features of Spanish accented Dutch might have affected the intelligibility of the Spanish speakers too, although this issue requires further investigation.

5.5 Conclusions

The results of the current study support all four predictions, and corroborate previous findings on the intelligibility of Spanish accented Dutch in showing that Dutch vowels pronounced by adult Spanish learners were frequently transcribed differently from their canonical forms by non-expert native Dutch listeners. These apparent vowel confusions are related to the native listeners' perceptual problems, which arise as a result of L1 constraints in the production of L2 vowels by Spanish learners. The numerous confusions, as evident from the high percentage of non-canonical transcriptions, point to the low intelligibility of separate Dutch words spoken by Spanish speakers, which could apparently hamper spoken interaction between Spanish learners and native listeners.

An additional goal of our study was to consolidate the usefulness of collecting transcriptions of L2 speech by selecting a diverse group of naïve native listeners. The consistency between the confusion patterns found in the present study and in our earlier crowdsource study (Burgos et al., 2015) demonstrates that involving the crowd's ear in speech research constitutes a promising approach yielding consistent data that can be employed to analyze the intelligibility of L2 speech and spotlight problematic areas of pronunciation.

In addition, we found evidence of perceptual adaptation on the part of the native listeners in our study, in that they were found to perceptually retune their native phoneme boundaries.

6 Matching acoustical properties and native perceptual assessments of L2 speech

This chapter has been reformatted and slightly modified from:

Burgos, P., Van Hout, R., & Planken, B. Matching acoustical properties and native perceptual assessments of L2 speech (submitted c).

6.1 Introduction

Learning the phonological system of a second language (L2) as an adult can be hard (Birdsong & Molis, 2001; Long, 1990), and many of the difficulties such learners encounter are related to interference from their native language (L1) (Cutler, 2012; Flege, Schirru, & MacKay, 2003). Several models, such as Flege's (1995) Speech Learning Model (SLM), have attempted to explain learners' difficulties in mastering the L2 phonological system in terms of the perceived similarity of segments in the L1 and L2. Flege and others agree that L2 contrasts based on fine-grained phonetic differences, particularly those contrasts that cover areas of the acoustic vowel space in which a single L1 native category is located, influence the perception and production of L2 phones, especially in the case of vowels (Baker & Trofimovich, 2005; Best, 1995; Bohn, 1995; Escudero, 2005; Flege, 1995; Flege et al., 2003; Major, 2001; McAllister, Flege, & Piske, 2002).

How can we assess whether L2 vowels are accurately produced? One way to investigate L2 vowel production accuracy is to conduct acoustical analyses. Comparing the acoustical properties of the target vowels produced by both L2 learners and native speakers can help establish whether the L2 realizations match those of the native speakers, and if not, where and to what degree there

is a mismatch (e.g., Guion, 2003; Iverson & Evans, 2007). However, an analysis of L2 vowel production based on acoustics alone does not automatically account for human vowel recognition, that is, how L2 vowels are perceived by native listeners. To determine the latter, we can ask native listeners to rate the intelligibility of L2 learners. Munro and Derwing (1995), define intelligibility as “the extent to which a speaker’s message is actually understood by a listener” (p. 76). Asking non-expert native listeners to orthographically transcribe L2 learners’ speech (Bent & Bradlow, 2003; Derwing & Munro, 1997) can help us establish whether a passage, sentence or word has been understood. Similarly, the intelligibility of vowels can be assessed by asking native listeners to transcribe separate, monosyllabic words containing target vowels, produced by L2 learners.

The aim of this article is to compare the acoustical properties of Dutch vowels produced by adult Spanish learners with the perception of these vowels by non-expert native Dutch listeners. Spanish learners of Dutch were asked to produce a series of monosyllabic Dutch words containing all the Dutch vowels (cf. Burgos, Jani, Van Hout, Cucchiarini, & Strik, submitted a for the acoustic mapping of the Spanish Dutch vowels onto the native Dutch vowels) and their vowel realizations were subsequently transcribed by non-expert native Dutch listeners to assess their intelligibility (cf. Burgos, Sanders, Van Hout, Cucchiarini, & Strik, submitted b for the outcomes with regard to perception). Comparing the acoustic properties of L2 vowel realizations with native listeners’ perceptions of these realizations can help determine which acoustic cues are present in the acoustic signal and which cues weigh most heavily in native listeners’ perception of L2 vowels.

A limited number of studies have investigated the relationship between L2 vowel production and native listeners’ vowel perception. Van Wijngaarden (2001) examined the perception of Dutch vowels embedded in CVC nonsense

words produced by native speakers of American English. Dutch vowels which had no equivalent in American English led to reduced recognition by the native listeners. Van Wijngaarden's (2001) findings show that vowels that are difficult for L2 subjects to produce are also difficult for native listeners to recognize. Munro (1993) studied the relationship between acoustical measurements (duration and spectrum) of English vowels produced by native speakers of Arabic and their corresponding accentedness ratings by linguistically trained native English listeners. His findings show that the native listeners rated the majority of L2 vowels as accented because they perceived durational and spectral deviances, most of which were attributed to specific characteristics of the Arabic vowel system. Munro (1993) also explored individual differences. With respect to duration, no clear patterns in the individual Arabic-English productions were observed, in terms of them being longer or shorter than the average duration in native productions. As to spectral properties, it was found that the Arabic subjects as a group did not differ significantly from the native English mean, whereas varying degrees of deviance from native English were found in their individual productions. In addition, Munro (1993) concluded that degree of accentedness did not correlate with learners' individual amount of experience, length of residence or daily use of English, suggesting that experience in the L2 does not guarantee success in achieving native-like pronunciation. Nevertheless, several studies have demonstrated that factors such as age of arrival, length of residence, formal instruction, amount of experience, L2 use and motivation play a role in L2 learning, particularly in the acquisition of L2 phones (cf. Moyer, 2013; Piske, MacKay, & Flege, 2001 for a review). These factors may differently affect the performance of individual learners who in turn may make use of different strategies in acquiring L2 vowels. This can result in large

variability in vowel realizations, both within and across learners with the same L1-L2 pairing (cf. Bent, Baesse-Berk, Borrie, & McKee, 2016).

Native listeners can rapidly perceive segmental deviations from the norm and can easily detect pronunciation errors by learners that native speakers are not likely to make (Magen, 1998). The variability inherent in L2 learners' production implies that native listeners have to adapt to different pronunciations across learners. For example, in the context of L2 vowel realizations, they have to be able to shift their category boundaries to accommodate an ambiguous vowel realization that differs from their usual expectations about phonemic categories (Cutler, 2012). These perceptual adaptation processes show that native listeners can adjust the L1 boundary between categories to accommodate variability in L2 speech, and benefit language processing (cf. Bradlow & Bent, 2008; Cutler, 2012).

The variability in L2 vowel realizations, and how such variability may be related to individual differences, has been understudied in language acquisition research. An additional aim of the present study is therefore to throw light on this potential relationship. Besides comparing the acoustical measurements of the Spanish learners' Dutch vowel productions with their corresponding perceptual assessments by native Dutch listeners, we will examine variability in production and perception in the context of the learners' proficiency level and additional factors that may play a role in L2 vowel accuracy, such as prior linguistic knowledge in multilinguals (De Angelis, 2007), length of residence and L2 use (cf. Piske et al., 2001).

Next, we consider the phonological properties of the Spanish and Dutch vowel systems. While Spanish has a straightforward five vowel system (/a, e, i, o, u/; see Figure 6.1) (Hualde, 2005), Dutch has a complex vowel system containing 15 full vowels (monophthongs: /i, y, u, I, ʏ, ɔ, ε, α, a:/; long mid vowels: /e:, ø:, o:/; diphthongs: /ei, œy, ou/; see Figure 6.1), next to the reduced

vowel /ə/ (Adank, Van Hout, & Smits, 2004b; Booij, 1995). Four features characterize further differences between the two vowel systems (see Table 6.1). Spanish does not have phonemic vowel length (Hualde, 2005; McAllister et al., 2002), whereas Dutch has a strict lax/tense distinction (lax vowels: /ɪ, ɛ, ə, ʏ, ʌ/; tense vowels: /i, y, u, e:, ø:, o:, a:/), which crosses the short/long distinction (short vowels: /i, y, u, ɪ, ʏ, ə, ɛ, ʌ/; long vowels: /a:, e:, ø:, o:, ei, œy, ou/) (Adank et al., 2004b). Spanish does not have the feature of front rounding, as all rounded vowels in Spanish are back vowels (/o, u/) (Hualde, 2005), while Dutch has four front rounded vowels (/ʏ, y, ø:, œy/). Spanish does not have diphthongs at the phoneme level, that is, single phonemes defined by their trajectory between two vowel positions; instead it has a rich inventory of 14 vowel combinations (Hualde, 2005). Dutch, on the other hand, has diphthongs at the phoneme level, such as /ei, œy, ou/. The Dutch long mid vowels (/e:, ø:, o:/) are not considered to be full diphthongs, but they are slightly diphthongized (cf. Adank et al., 2004b; Van der Harst, Van de Velde, & Van Hout, 2014). Finally, Spanish is distinguished by three height values (high: /i, u/; mid: /e, o/; low: /a/) (Hualde, 2005), whereas Dutch is characterized by four height values (high: /i, y, u/; high mid: /ɪ, ʏ, e:, ø:, o:/; low mid: /ɔ, ɛ, ei, œy, ou /; low: /ʌ, a:/) (Booij, 1995), as shown in Figure 6.1.

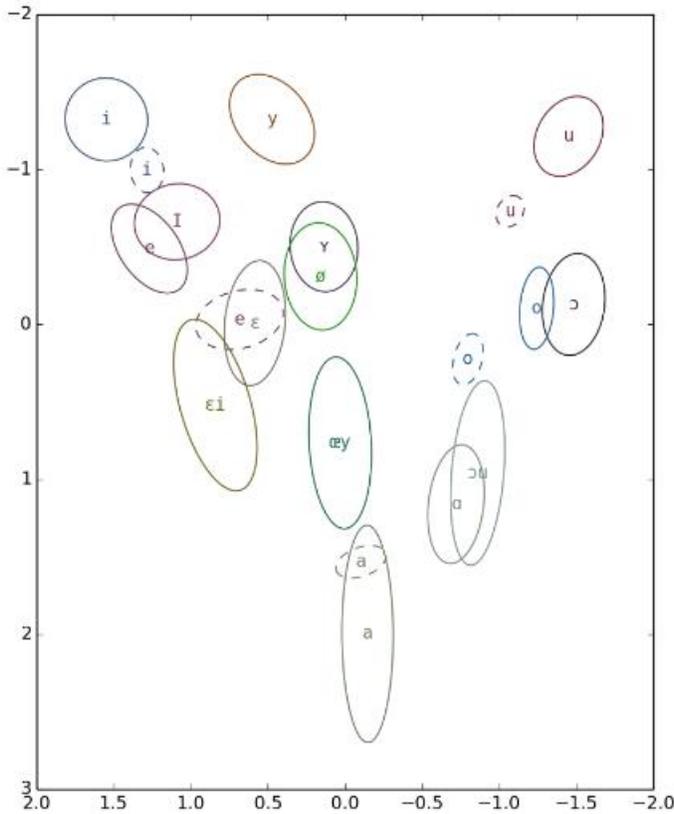


Figure 6.1 $F1$ (y-axis) and $F2$ (x-axis) (normalized) values for the five Spanish vowels (dashed lines), as realized by Spanish L1 speakers and for the fifteen Dutch vowels, as realized by Dutch L1 speakers; all vowels are measured at 50% of the vowel duration; the mean values are indicated by the vowel symbols; the values for Spanish vowels were drawn from Chládková, Escudero, and Boersma (2011); the values for Dutch values were drawn from Van der Harst (2011).

Table 6.1 *Distinctive features of the Spanish and Dutch vowel systems.*

	Spanish	Dutch	Comment
Length (duration)	–	+	Dutch has a complex interaction between the distinctions of lax/tense and short/long.
Front rounding	–	+	Both languages have back rounded vowels.
Diphthong	–	+	Spanish has 14 vowel combinations, but no diphthongs at the phoneme level.
Height	3 levels	4 levels	Dutch has an additional height value, as can be observed in the vowel space of Figure 6.1.

Previous studies on the acoustical properties of Dutch vowels produced by adult Spanish learners of Dutch (cf. Burgos et al., submitted a) and on the perceptual assessments of the same vowels by non-expert native Dutch listeners (cf. Burgos et al., submitted b) have shown that Spanish learners have problems with contrasts in vowel height, vowel length, front rounding and diphthongization. Considering previous findings and given the different types of vowel confusions observed in both production and native perception, we predict that the acoustic cues (vowel height, length, rounding and diphthongization) in the speech signal of the learners' vowel realizations will have a direct impact on both production and perception outcomes. We expect that these acoustic cues may have different weightings in learners' production and, particularly, in native perception, and that they will affect the variability in the production and perception confusion patterns.

6.2. Method

6.2.1 Spanish learners

The speech of 28 adult Spanish learners of Dutch (9 males and 19 females), originating from Spain and a number of Latin American countries (Argentina, Dominican Republic, Guatemala, Mexico and Venezuela), was employed as

stimulus material. All the learners were highly educated and living in the Netherlands at the time of the study. They had already followed or were taking Dutch courses and all of them reported using Dutch in daily life. As all the learners were familiar with the Common European Framework of Reference for Languages (CEFR; Council of Europe), they were asked to assess their own proficiency level in Dutch, and in other foreign languages they spoke, using the CEFR Self-Assessment Grid (Council of Europe). They rated themselves in Dutch at one of the following four CEFR levels: A1 (CES (=Cambridge English Scales) 100-119), A2 (CES 120-139), B1 (CES 140-159) and B2 (CES 160-180) (UCLES, 2015). Table 6.2 shows information about the Spanish learners per CEFR language proficiency level (cf. Burgos et al., submitted a, an earlier acoustic study, for more detailed information about the Spanish learners).

Table 6.2 Means and standard deviations (SD) for speaker variables per CEFR language proficiency level; Age = age at the time of the recording, in years, AoA = age of arrival in the Netherlands, in years, LoR = length of residence in the Netherlands, in years, Use of Dutch = Self-estimated daily use of Dutch, in hours, Origin = where speakers were born and brought up.

CEFR	Age	AoA	LoR	Use of Dutch	Origin	
					Spain	Latin America
A1	32.2 (7.5)	30.2 (6.6)	2.0 (3.1)	6.7 (3.5)	8	2
A2	31.6 (5.3)	30.6 (5.5)	0.9 (0.4)	7.8 (2.5)	6	1
B1	40.2 (1.2)	32.2 (3.3)	8.0 (3.9)	5.0 (2.0)	2	2
B2	41.0 (7.8)	30.8 (7.8)	10.1 (5.9)	13.1 (8.0)	5	2
Overall	35.4 (7.6)	30.8 (6.0)	4.6 (5.3)	8.3 (5.4)	21	7
Range	20–52	19.3–42.0	0.1–20.0	2–24		

6.2.2 Stimulus materials

The stimulus materials in the present study are from an existing corpus of Spanish L1 Dutch L2 (cf. Burgos, Jani, Cucchiarini, Van Hout, & Strik, 2014b) which includes systematic productions of Dutch sounds that are problematic for Spanish learners. The material we used in the present study comprises a list of Dutch monosyllabic words read out by adult Spanish learners. The same material was used by Van der Harst (2011), and Van der Harst et al. (2014), who obtained recordings of the same list of Dutch words produced by native Dutch speakers. The word list used in the original corpus comprises 278 monosyllabic and disyllabic words representing all the Dutch vowels in different contexts. For our study, we employed a subset of these 278 words, namely 29 Dutch monosyllabic words per speaker. This subset of 29

Dutch words included all 15 Dutch vowels in stressed position followed either by /s/ or /t/, as vowel quality is known to alter only minimally when the vowel is followed by these consonantal sounds (Van der Harst, 2011, p. 146; Van der Harst et al., 2014, p. 254). Table 6.3 provides an overview of the 29 Dutch words containing all 15 Dutch vowels, and their corresponding phonological and orthographic representations. No example of a word containing the vowel /y/ followed by /s/ was included, as this combination does not exist in Dutch monosyllabic nouns.

Table 6.3 Selected *-s* and *-t* words as speech stimuli (Van der Harst, 2011); Phon = phonological representation (in IPA), Orth = orthographic representation.

Vowel	Phon	Orth	s-word	t-word		
Monophthongs	/i/	<ie>	<i>kies</i>	/kis/	<i>riet</i>	/rit/
	/y/	<uu>	-	/y/	<i>fuut</i>	/fyt/
	/u/	<oe>	<i>poes</i>	/pus/	<i>voet</i>	/vut/
	/ʌ/	<i>	<i>vis</i>	/vɪs/	<i>fit</i>	/fɪt/
	/ʏ/	<u>	<i>zus</i>	/zʏs/	<i>put</i>	/pʏt/
	/ɔ/	<o>	<i>vos</i>	/vɔs/	<i>vlot</i>	/vlot/
	/ɛ/	<e>	<i>zes</i>	/zɛs/	<i>vet</i>	/vɛt/
	/ɑ/	<a>	<i>gas</i>	/xɑs/	<i>rat</i>	/rat/
	/a:/	<aa>	<i>aas</i>	/a:s/	<i>staat</i>	/sta:t/
Long mid vowels	/e:/	<ee>	<i>mees</i>	/me:s/	<i>beet</i>	/be:t/
	/ø:/	<eu>	<i>neus</i>	/nø:s/	<i>neut</i>	/nø:t/
	/o:/	<oo>	<i>boos</i>	/bo:s/	<i>boot</i>	/bo:t/
Diphthongs	/ɛi/	<ij>	<i>ijs</i>	/ɛis/	<i>spijt</i>	/speit/
	/œy/	<ui>	<i>huis</i>	/hœys/	<i>fluit</i>	/flœyt/
	/ou/	<ou>	<i>kous</i>	/kous/	<i>fout</i>	/fɔut/

All non-native tokens were recorded in a quiet room at the Linguistics Department of the Radboud University Nijmegen or at the speakers' home, using a headset (Logitech, USB entry DZL-A-0001 4-B) and a laptop (ACER AMD Quad-core Processor A6-3400M with Turbo CORE Technology up to 2.30 GHz). The data were recorded at a sampling frequency of 16 kHz. The Spanish learners read out loud the Dutch words, which were presented on a computer screen one by one, at intervals of three seconds. Each word from the set of 29 words was recorded by each speaker only once, which resulted in 812 word tokens (29 words x 28 speakers). Six speech samples (from six different subjects) were discarded due to erroneous recording. Thus, a total of 806 word tokens from the Spanish learners were subjected to analysis.

We used the same set of 29 words spoken by 20 native speakers of Standard Dutch (10 males and 10 females), collected by Van der Harst (2011) and Van der Harst et al. (2014) to describe the Dutch vowel system. The Dutch samples were analyzed (see below) using similar techniques as those employed for the non-native data. Specific details about the native Standard Dutch corpus are reported in Van der Harst (2011).

6.2.3 Analysis of the speech recordings

The words read by the native Dutch speakers and by the Spanish learners were orthographically transcribed in Praat (Boersma & Weenink, 2010), and subsequently segmented following the procedures for vowel segmentation described in Van Son, Binnenpoorte, Van den Heuvel, and Pols (2001), and Van der Harst (2011). 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 done by an experienced transcriber, the first author, and was then checked by a native Dutch phonetician.

6.2.4 *Acoustical analyses*

Acoustical analyses were performed to extract measurements of the first and second formants (F1 and F2) and of the duration of the Dutch vowels produced by adult Spanish learners. It should be noted that measurements of third formant (F3) were not analyzed in the current study, as previous research has shown that F3 is not essential in the identification of front rounded vowels in Dutch, and that using F1 and F2 only is sufficient to identify these vowels (Adank, 2003; Cohen, Slis, & 't Hart, 1963,1967; Van der Harst, 2011). The first two formants were measured at three equidistant points (i.e., at 25%, 50% and 75% of the vowel duration). This information helps to determine if and how diphthongization is realized by Spanish learners producing all mid vowels and diphthongs in comparison to native speakers, as mid vowels and diphthongs in Dutch are long and show a milder or stronger degree of diphthongization (Adank et al., 2004b; Van der Harst et al., 2014). All measurements were automatically extracted using an LPC (Linear Predictive Coding) analysis. First, every vowel token was assigned a specific number of coefficients, i.e., four, five or six coefficients. Next, an LPC script based on the chosen number of coefficients was run in order to extract F1 and F2 values. The same procedure was repeated for the measurement extractions of duration. All resulting measurements were manually checked by the first author who corrected any errors found. Subsequently, an additional check for outliers, checked at 25%, 50% and 75%, was carried out by the same native Dutch phonetician mentioned earlier, who followed the procedure employed in Van der Harst (2011) and Van der Harst et al. (2014), and corrected any errors found. All vowel realizations were then normalized using Lobanov's (1971) transformation to neutralize formant frequency variations resulting from anatomic differences among informants (cf. Adank, Smits, & Van Hout, 2004a). Durational values were also normalized. For a detailed description of

the non-native speech data (i.e., stimulus materials, analysis of the speech recordings and acoustic analyses) used in the present study, see Burgos et al. (submitted a).

6.2.5 Native Dutch listeners

A snowball sampling strategy had been employed to recruit native listeners in an earlier study on the perception of Spanish-Dutch speech (Burgos et al., submitted b). This sampling technique consists of recruiting subjects from the social networks of a starting set of individuals. In the earlier study, each individual in the starting set was asked to recruit at least five native speaker subjects from his/her networks of family and friends, in order to reach a heterogeneous group of native listeners. The individuals themselves could not take part in the experiment. The same sampling technique was used in the present study. The starting set of 25 individuals (7 males and 18 females) in the present study were all undergraduate students of International Business Communication (IBC, Department of Communication and Information Studies) at the Radboud University Nijmegen, in the Netherlands. The native listeners they recruited had to meet the following criteria: 1) at least 18 years old, 2) native Dutch speaker, 3) not linguistically trained, and 4) unfamiliar with Spanish-accented Dutch. A total of 139 native Dutch listeners who met the criteria were recruited. They were all asked to participate in a transcription task, which required them to transcribe the non-native (Spanish-Dutch) stimuli. A total of 132 native Dutch listeners (59 males and 73 females) completed the transcription task. The transcriptions of seven listeners who did not complete the task were discarded. The listeners were heterogeneous in terms of age (range: 18-66 yr old, $M = 32.39$, $SD = 16.26$) and completed education (elementary school ($n = 4$), high school ($n = 73$), vocational training education ($n = 24$), higher professional education ($n = 24$), university degree

($n = 7$). It should be noted that the transcription task consisted of the non-native tokens interspersed with native tokens of two (one male and one female) from the corpus of native speakers of Standard Dutch mentioned earlier. Native stimuli were included to increase the validity of the task, as these native tokens could be used by the transcribers as anchor points (cf. Burgos et al., submitted b). The transcriptions of both the non-native and native tokens of all 132 listeners were used in subsequent analyses.

6.3 Results

Subsection 6.3.1 presents the results of the acoustic measurements of the non-native and native data. Subsection 6.3.2 focuses on the native listeners' perceptions of Dutch vowels produced by the adult Spanish learners and the native Dutch speakers. The final subsection, 6.3.3, compares the outcomes of the production study (acoustic data) with those of the perception study (listener data) (cf. Burgos et al., submitted b).

6.3.1 Acoustic data

The Dutch vowels produced by Spanish learners and native Dutch speakers were analyzed using a multinomial logistic regression, which is a statistical classification technique. It is used to predict a vowel classification using a categorically distributed dependent variable, given a set of predictor variables. Based on the acoustic values of the non-native and native speech data, the regression calculates the probabilities of canonical (target vowel) and non-canonical classifications of the vowel realizations. A given vowel realization is classified in the vowel category with the highest probability.

We investigated the non-native and native data using three classification conditions to determine to what degree outcomes depend on the vowel sets to be classified, namely “Total”, “Group” and “Individual”. In the classification

condition “Total”, we pooled the non-native and native data, in “Group”, the non-native and native data were treated as two independent sets or groups, and in “Individual”, the individual Spanish learner data were added, learner by learner, to the data of the native group. The regressions for each classification condition were conducted using F1 and F2 only, and F1, F2 plus vocalic duration. In this way, the analyses could throw light on the extent to which duration plays a role in vowel classification.

6.3.1.1 Non-native data

We first focus on the results in the three classification conditions. Subsequently, we present the non-native matrix that we obtained in the condition “Group” using F1, F2 and duration.

Table 6.4 presents the average percentage of canonical classifications for the three conditions “Total”, “Group” and “Individual”, with F1 and F2 (at 25%, 50%, 75% of the vowel duration) only, and with F1, F2 and duration. It can be seen that the average percentage of canonical classifications in the condition “Total” with F1, F2 only, at 61.1%, increases to 72.5% after including duration. A similar increase can be seen in “Group” (61.7% → 74.7%) and “Individual” (72.4% → 89.1%).

However, an increase in the average percentage of canonical classifications does not necessarily mean that each target vowel benefits equally from including duration. Upon closer inspection, we see that in the “Total” condition, almost all target vowels benefit from duration, and adding duration is not detrimental to any target vowel. But three target vowels, the diphthongs <ij>, <ui> and <ou>, are not affected by adding duration; the percentages of canonical classifications for these vowels remain unchanged. Along similar lines, in the “Group” condition, almost all target vowels benefit from duration, but the canonical classification percentages for the target vowels <e> and <ui>

are not affected by duration. Finally, in the “Individual” condition, the canonical transcription percentages for 13 target vowels increase by adding duration, but those for <ie> and <e> are not affected. In sum, this indicates that with duration, the average percentage of canonical classifications is consistently higher in each classification condition, and for the majority of target vowels. This improvement indicates that duration contributes substantially to an increase in the probability of a canonical classification.

Table 6.4 Means and standard deviations (SD) for the canonical classifications of the Dutch vowels produced by Spanish learners for the three classification conditions “Total”, “Group” and “Individual”, with both F1 and F2 (25%, 50%, 75%) only, and F1, F2 and duration, and target vowels whose percentages (do not) change after including duration, deviations $\leq 2.5\%$ after including duration are rated as unchanged; ClassificationC = classification condition, + dur = including duration, Vowel + dur \uparrow = target vowels whose percentages of canonical classifications improve after including duration, Vowel + dur \downarrow = target vowels whose percentages of canonical classifications do not improve after including duration, Vowel + dur \leftrightarrow = target vowels whose percentages of canonical classifications remain unchanged after including duration.

ClassificationC	F1, F2	F1, F2 + dur	Vowel + dur \uparrow	Vowel + dur \downarrow	Vowel + dur \leftrightarrow
Total	61.1 (15.0)	72.5 (13.8)	<ie>,<uu>,<oe>,<i>,<u>,<o>,<e>, <a>,<aa>,<ee>,<eu>,<oo>	-	<ij>,<ui>,<ou>
Group	61.7 (15.0)	74.7 (13.3)	<ie>,<uu>,<oe>,<i>,<u>,<o>,<a>, <aa>,<ee>,<eu>,<oo>,<ij>,<ou>	-	<e>,<ui>
Individual	72.4 (17.9)	89.1 (13.2)	<uu>,<oe>,<i>,<u>,<o>,<a>,<aa>, <ee>,<eu>,<oo>,<ij>,<ui>,<ou>	-	<ie>,<e>

An improvement (see Table 6.4) is also observed across the three conditions (“Total” \rightarrow “Group” \rightarrow “Individual”). The average percentage of canonical classifications using F1 and F2 slightly improves from “Total” (61.1%) to “Group” (61.7%), and a substantial improvement is observed from

“Group” to “Individual” (72.4%). A similar improvement pattern from “Total” (72.5%) to “Group” (74.7%) to “Individual” (89.1%) is found when duration is added (F1, F2 + duration).

The inclusion of native speakers of Dutch in the “Total” condition does not appear to offer additional resources for classifying the Spanish vowel realizations. The improvement to classifications in the “Group” and “Total” conditions are comparable. The “Individual” condition gives a boost to the canonical scores. In this condition, the vowels of one Spanish learner are classified amongst the vowels of all 20 native speakers. The “Individual” condition can only produce an improvement if the Spanish learners actually produce distinct acoustic properties between the target vowels, however weak. Adding more Spanish learners seems to blur these already vulnerable distinctions, and this effect is probably augmented by the fact that the distinctions are highly variable across and within learners.

Is this improvement in the average percentage of canonical classifications across classification conditions – after including duration – also found for the canonical classifications per target vowel? Table 6.5 shows the degree of improvement in the percentages of canonical classifications per target vowel and the average percentage of canonical classifications per condition (“Total” → “Group” → “Individual”), using F1 F2 and duration, for all 15 Dutch vowels produced by Spanish learners.

Table 6.5 Degree of improvement for the percentages of canonical classifications per target vowel and per classification condition (“Total” → “Group” → “Individual”) for all 15 Dutch vowels produced by Spanish learners (F1, F2 + duration); Vowel = target vowel, Δ T-G = difference between the percentages of canonical classifications of the classification conditions “Total” and “Group”, Δ G-I = difference between the percentages of canonical classifications of the classification conditions “Group” and “Individual”, %Can = average percentage of canonical classifications.

Vowel	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	%Can
Total	62.5	53.6	76.4	69.6	66.1	76.8	89.1	85.7	78.6	83.6	81.8	85.7	76.8	38.2	63.6	72.5
Group	67.9	53.6	72.7	76.8	71.4	83.9	90.9	83.9	83.9	85.5	78.2	83.9	82.1	41.8	63.6	74.7
Δ T-G	5.4	-	-3.7	7.2	5.3	7.1	1.8	-1.8	5.3	1.9	-3.6	-1.8	5.3	3.6	-	2.2
Individual	69.6	92.9	94.5	50.0	83.9	89.3	96.4	94.6	98.2	96.4	98.2	96.4	94.6	85.5	96.4	89.1
Δ G-I	1.7	39.3	18.1	-26.8	12.5	5.4	5.5	10.7	14.3	10.9	20.0	12.5	12.5	43.7	32.8	14.4

The average percentage of canonical classifications in “Total” (72.5%) is similar to that in “Group” (74.7%), showing only a modest improvement of 2.2%. There are no large differences between the target vowels, but what differences there are clearly indicate that classification depends on the properties of the data to be classified. A large improvement (14.4%) is observed from “Group” (74.7%) to “Individual” (89.1%), with large differences between the individual target vowels also. The rounded target vowels, in particular, seem to benefit (Δ <uu> = 39.3%, Δ <u> = 12.5%, Δ <eu> = 20.0%, Δ <ui> = 43.7%), whereas the outcome for <i> is the only negative pattern (Δ = -26.8%), which may indicate that the distinction <i> versus <ie> is hard to classify at the level of individual Spanish learners. The striking improvement in the average percentage of canonical classifications from “Group” to “Individual” (14.4%), as well as in the percentages of canonical classifications per target vowel (with the exception of <i>) shows

that the individual Spanish learners are making distinctions between the Dutch vowels, however weak and variable these may be.

We now focus on the outcomes of the non-native acoustic data in the classification condition “Group”, as this is the condition, unlike “Total”, in which the non-native data is treated as an independent group. Therefore, the outcomes in this condition can best classify the Spanish-accented Dutch vowels based on their own characteristics. We present the results in the “Group” condition employing F1, F2 and duration, as formant frequencies (F1 and F2) and duration are both important properties to be taken into account when characterizing Dutch vowels.

Table 6.6 presents the matrix representing the classification of the Spanish learners’ Dutch vowel realizations, using a multinomial logistic regression. The columns represent the 15 target vowels corresponding to the nine monophthongs (<ie>, <uu>, <oe>, <i>, <u>, <o>, <e>, <a>, <aa>), the three long mid vowels (<ee>, <eu>, <oo>), and the three diphthongs (<ij>, <ui>, <ou>). The rows show the overall percentages of canonical (marked green) and non-canonical classifications for the 15 target vowels. The column Total shows the average percentage of the sum of all percentages of classified vowels per row.

Table 6.6 *Probability ratio of canonical (indicated in green) and non-canonical classifications for all 15 Dutch vowels produced by Spanish learners in the classification condition “Group” (F1, F2 + duration); target vowels in the columns, classified vowels in the rows, non-canonical classifications with deviations > 2.5% related to vowel height (in pink), vowel length (in turquoise), rounding (in yellow) and diphthongization (in teal) are also indicated; Vow = target vowel.*

Vow	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	Total
ie	67.9	0.0	0.0	14.3	1.8	0.0	0.0	0.0	0.0	9.1	0.0	0.0	0.0	1.8	0.0	6.6
uu	0.0	53.6	3.6	0.0	5.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.6	0.0	2.7
oe	0.0	21.4	72.7	0.0	14.3	3.6	0.0	0.0	0.0	0.0	0.0	5.4	0.0	1.8	3.6	7.7
i	21.4	0.0	0.0	76.8	5.4	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.3
u	1.8	14.3	9.1	3.6	71.4	3.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	6.8
o	0.0	0.0	3.6	0.0	1.8	83.9	0.0	1.8	0.0	0.0	0.0	3.6	0.0	0.0	3.6	6.8
e	0.0	0.0	0.0	5.4	0.0	0.0	90.9	1.8	0.0	3.6	3.6	0.0	3.6	1.8	0.0	7.6
a	0.0	0.0	0.0	0.0	0.0	0.0	0.0	83.9	14.3	0.0	0.0	0.0	0.0	5.5	1.8	7.3
aa	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.1	83.9	0.0	0.0	0.0	0.0	0.0	5.5	6.7
ee	3.6	0.0	0.0	0.0	0.0	0.0	3.6	0.0	0.0	85.5	0.0	0.0	8.9	0.0	0.0	6.9
eu	0.0	3.6	1.8	0.0	0.0	0.0	1.8	0.0	0.0	0.0	78.2	0.0	1.8	12.7	0.0	6.7
oo	0.0	7.1	7.3	0.0	0.0	8.9	0.0	0.0	0.0	0.0	83.9	0.0	0.0	0.0	7.3	7.7
ij	5.4	0.0	0.0	0.0	0.0	0.0	1.8	1.8	0.0	1.8	0.0	0.0	82.1	9.1	0.0	7.1
ui	0.0	0.0	1.8	0.0	0.0	0.0	0.0	3.6	1.8	0.0	16.4	0.0	3.6	41.8	12.7	5.6
ou	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	7.1	0.0	21.8	63.6	6.5

The outcomes show that 74.7% of all classifications are canonical, whereas 25.3% are non-canonical. The highest percentage of canonical classifications is found for the vowel <e> (90.9%), and the lowest for <ui> (41.8%).

The variety in non-canonical classifications indicates that the Dutch vowels produced by the learners were classified on the basis of their acoustic properties as different vowels than the target vowels. The highest differentiation was found for the target diphthongs <ui> and <ou>, classified as eight and seven different non-canonical vowel categories respectively, whereas the lowest differentiation was found for the target vowel <aa>, assigned to only two different vowel categories, followed by the target vowels <i>, <o>, <ee>, <eu>, and <oo>, with three categories each.

Upon closer inspection, Table 6.6 shows non-canonical classifications related to problems with vowel height (e.g., in the vowel confusion <i>-<ie>), vowel length (e.g., in the confusions <a>-<aa> and <o>-<oo>), rounding (e.g., in the confusions <uu>-<u>-<oe> and <ui>-<ou>), and diphthongization (e.g., in the confusions <eu>-<ui> and <oo>-<ou>). Clearly, these problems relate to the four distinctive features listed in Table 6.1. An examination of the most conspicuous vowel confusions with values higher than 10% brings us to the pair <i>-<ie>. The target vowel <ie> (67.9%) is extensively classified as <i>, as attested by the high percentage of non-canonical classifications, namely 21.4%. In contrast, the target vowel <i> (76.8%) is less frequently classified as <ie> (14.3%). Asymmetry is also found in confusions related to vowel length, as is the case for the vowel pair <a>-<aa>. The target vowel <aa> (83.9%) is often classified as <a> (14.3%), more than <a> is as <aa> (7.1%). A similar situation applies to the target vowel <o> (83.9%), which is frequently classified as non-canonical <oo> (8.9%). Vowel confusions related to rounding are found for the target front rounded vowels <uu> (53.6%) and <u> (71.4%), which are frequently confused with each other, and, especially, with the back rounded vowel <oe>, yielding non-canonical percentages of 21.4% and 14.3% respectively. Rounding is also involved in the <ui>-<ou> confusion. The target front rounded vowel <ui> (41.8%) is often classified as <ou>, and is given the highest percentage of non-canonical classifications in the non-native matrix (21.8%), a higher percentage than for the target back rounded vowel <ou> (63.6%) which is classified as <ui> (12.7%). As to diphthongization, the target long mid vowel <eu> (78.2%) is frequently classified as <ui> (16.4%), more than <ui> is classified as <eu> (12.7%). This indicates that the long mid vowel <eu> is extremely diphthongized by the Spanish learners.

6.3.1.2 Native data

The Dutch vowels produced by native Dutch speakers were analyzed in the classification conditions “Total”, in which the native and non-native data are pooled together, and in the classification condition “Group”, in which the native and non-native data are treated as two independent groups, using F1 and F2 only, and F1, F2 and duration.

Table 6.7 shows the average percentages of canonical classifications for the two classification conditions “Total” and “Group”, with both F1 and F2 (at 25%, 50%, 75% of the vowel duration) only, and F1, F2 and duration. The average percentage of canonical classifications in “Total” with F1, F2 only, at 84.3%, increases to 91.3% after including duration. A similar increase after adding duration is seen in “Group” (95.2% → 99.2%). In “Total”, ten target vowels benefit from duration, whereas <ie> and <uu> show a decrease. Three target vowels, <oe>, <ij> and <ui> (with deviations $\leq 2.5\%$), are not affected by adding duration. In “Group”, six target vowels benefit from duration, while the other nine target vowels do not. Overall, including duration leads to consistently higher percentages of canonical classifications per target vowel, although this seems to be less beneficial for the native vowel classifications than it is for the non-native vowel classifications.

Table 6.7 Means and standard deviations (SD) for the canonical classifications of the Dutch vowels produced by native Dutch speakers for the two classification conditions “Total” and “Group”, with both F1 and F2 (25%, 50%, 75%) only, and F1, F2 and duration, and target vowels whose percentages (do not) change after including duration, deviations $\leq 2.5\%$ after including duration are rated as unchanged; ClassificationC = classification condition, + dur = including duration, Vowel + dur \uparrow = target vowels whose percentages of canonical classifications improve after including duration, Vowel + dur \downarrow = target vowels whose percentages of canonical classifications do not improve after including duration, Vowel + dur \leftrightarrow = target vowels whose percentages of canonical classifications remain unchanged after including duration.

ClassificationC	F1, F2	F1, F2 + dur	Vowel + dur \uparrow	Vowel + dur \downarrow	Vowel + dur \leftrightarrow
Total	84.3 (11.7)	91.3 (8.3)	<i>,<u>,<o>,<e>,<a>,<aa>, <ee>,<eu>,<oo>,<ou>	<ie>,<uu>	<oe>,<ij>,<ui>
Group	95.2 (6.2)	99.2 (2.2)	<i>,<u>,<o>,<ee>,<eu>, <oo>	-	<ie>,<uu>,<oe>,<e>,<a>, <aa>,<ij>,<ui>,<ou>

The average percentage of canonical classifications of the native data using F1 and F2 only improves from “Total” (84.3%) to “Group” (95.2%), with a greater improvement (10.9%) than that found for “Total” \rightarrow “Group” in the non-native data (i.e., 2.1%; see Table 6.5). A similar improvement to that in the non-native data is found when duration is included (F1, F2 + duration): “Total” (91.3%) \rightarrow “Group” (99.2%). This substantial improvement seems to suggest that the presence of non-native data in “Total” detrimentally affects the outcomes of the native speech data, indicating that the statistical classifier adapts its classification when non-native data are included.

To understand this mechanism better, Table 6.8 shows the native matrix in the classification condition “Total” with F1, F2 and duration. Most non-canonical classifications seem to be related to the vowel confusions observed in the non-native matrix, for example, to the vowel confusions in the pairs

<ie>-<i> and <i>-<ie>, associated with vowel height. The <uu>-<u> confusion is related to height as well. Other confusion patterns, associated with rounding, are observed too, such as the non-canonical classifications of the target front rounded <uu> as the front unrounded <i>, or of the target back unrounded <a> as the back rounded <o>. Diphthongal confusions are found for <eu> and <oo>. The feature that does not reflect problems in this matrix is length, indicating that duration does not lead to confusions in the native data because it is a secondary feature in the native-produced vowel distinctions. These outcomes seem to suggest that problematic features found in the statistical classifications of the non-native vowels recur in the classifications of the native vowels when these features are of primary relevance in the native-produced vowel distinctions. The classifier would seem to have adapted to the great variability in the vowel realizations in the non-native data with detrimental results for the native data as a consequence (cf. Berck (2017) who shows that machine learning algorithms are affected by infusing errors in linguistic data).

Table 6.8 *Probability ratio of canonical (indicated in green) and non-canonical classifications for all 15 Dutch vowels produced by native Dutch speakers in the classification condition “Total” (F1, F2 + duration); target vowels in the columns, classified vowels in the rows, non-canonical classifications with deviations > 2.5% related to vowel height (in pink), vowel length (in turquoise), rounding (in yellow) and diphthongization (in teal) are also indicated; Vow = target vowel.*

Vow	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	Total
ie	85.0	0.0	0.0	17.5	0.0	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0	0.0	7.2
uu	2.5	75.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0	2.9
oe	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.9
i	12.5	5.0	0.0	77.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.4
u	0.0	20.0	0.0	0.0	95.0	0.0	7.5	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0	7.9
o	0.0	0.0	0.0	0.0	0.0	97.5	0.0	7.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.2
e	0.0	0.0	0.0	2.5	5.0	0.0	90.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0	6.9
a	0.0	0.0	0.0	0.0	0.0	0.0	0.0	92.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.4
aa	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	2.5	7.1
ee	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0	0.0	95.0	0.0	0.0	7.5	0.0	0.0	7.2
eu	0.0	0.0	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0	80.0	0.0	2.5	0.0	0.0	5.9
oo	0.0	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0	0.0	0.0	95.0	0.0	0.0	0.0	6.7
ij	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	0.0	0.0	90.0	0.0	0.0	6.4
ui	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.5	0.0	0.0	100.0	0.0	7.8
ou	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	97.5	7.1

6.3.2 Listener data

In this subsection, we focus on human vowel recognition, by examining the native Dutch listeners’ transcriptions of the Dutch vowels produced by adult Spanish learners (non-native matrix) and by native Dutch speakers (native matrix) (cf. Burgos et al., submitted b for a detailed description of the perception outcomes).

It should be noted that the non-native and native matrices consist of 15 columns representing the 15 target vowels x 15 rows representing the classified vowels. This was not the case for the non-native and native matrices of the listener data in the earlier perception study (cf. Burgos et al., submitted b), which contained 15 columns representing the 15 target vowels x 20 rows

representing the transcribed vowels. That is, the 20 rows consisted of 15 rows for the 15 Dutch vowels and five additional rows: one row for the frequent non-canonical variant <ai> (overdiphthongized vowel combination assigned to the target vowel <ij>), and four rows in the Rest category (containing transcribed vowels < 5%), including non-canonical variants related to longer duration, diphthongization, other transcriptions and consonants. With a view to comparing the non-native and native matrices of the acoustic data and the non-native and native matrices of the listener data, we decided to alter the number of rows in the original matrices of the listener data. We subsumed the <ai> transcriptions under <ij>, and, subsequently, distributed the percentages of the four rows in the Rest category throughout the remaining 15 rows representing the transcribed vowels. This allowed us to compare the two 15x15 matrices for the non-native (see Table 6.9) and native data (see Table 6.10).

6.3.2.1 *Non-native data*

Table 6.9 presents the non-expert native Dutch listeners' transcriptions of the Dutch vowels produced by the 28 adult Spanish learners. The rows show the overall percentages of canonical (indicated in green) and non-canonical transcriptions of the 15 target vowels.

Our outcomes show that 65.4% of all classifications are canonical, whereas 34.6% are non-canonical. The highest percentage of canonical transcriptions was for the vowel <e> (87.3%), while the lowest was for the <uu> (33.8%). The last two vowels in the list, namely, the front rounded <u> and <uu>, do not occur in Spanish and can be considered new for Spanish learners (cf. Flege, 1995).

The most striking difference in comparing Table 6.9 to Table 6.6 (the acoustic data), is that the non-native matrix of the listener data is much more

distributed, in that it shows more variability in the vowel confusions. This indicates that the listeners perceived the vowels spoken by the Spanish learners in various ways. The highest variability was found for <ui> and <eu>, assigned to 14 and 13 different non-canonical vowel categories respectively. The degree of variability here might be related to the fact that <ui> and <eu> are front rounded vowels. The lowest variability is found for the target vowel <i>, even though it is assigned to as many as six different non-canonical vowel categories.

The non-canonical transcriptions in Table 6.9 clearly show what the vowel confusions are, and therefore, which features were perceived by the native listeners to be (erroneously) employed by the Spanish learners when producing the Dutch target vowels. The most outspoken vowel confusion, related to vowel height, is observed for the target vowel <i> which is assigned to <ie> (44.2%). Non-canonical transcriptions of the target vowel <ie> as <i> (13.7%) were also found, but at a substantially lower percentage, indicating that there is an asymmetrical confusion between these vowels. Other confusions related to vowel height were observed in the target vowel <u>, which was frequently transcribed as <uu> (10.2%), more than <uu> as <u> (4.0%). The target diphthong <ou> was often perceived as <oo> (19.0%), although <oo> was seldom transcribed as <ou> (1.8%). The most conspicuous confusions related to vowel length were found in the short monophthongs <uu>, <i>, <o> and <a>, which were often perceived as having longer duration, and in the long vowel <aa> and the long mid vowels <ee>, <eu> and <oo>, which were perceived as monophthongs with shorter duration. In other words, the target short monophthongs were often perceived as long vowels, and the target long vowels as short vowels. An asymmetrical confusion related to vowel length can be seen in the vowels in the pair <a>-<aa>, which are hard

to distinguish, with <a> more often perceived as <aa> (30.8%) than <aa> as <a> (14.2%).

With respect to the front rounded vowels, Table 6.9 shows that the front rounded monophthongs <uu> and <u> are perceived as the back rounded vowel <oe> (30.3% and 37.0% respectively). A similar pattern was found for the front rounded diphthong <ui> which was often transcribed as the back rounded diphthong <ou> (14.6%). However, the front rounded long mid vowel <eu> was perceived differently: either as a front vowel (<ij>, 3.9%) or as a back rounded vowel (<oo>, 9.6%). We found fewer vowel confusions related to diphthongization than to vowel height, vowel length and rounding. The target long mid vowels <ee> and <eu> were often perceived as <ij> and <ui> respectively, indicating that they were overdiphthongized. It should be noted that the results reported in Burgos et al. (submitted b) also show that native Dutch listeners perceived the extreme diphthongization with which some Dutch vowels were produced by the Spanish learners, especially in the case of the long mid vowels and diphthongs. Finally, the frequent assignment of the non-canonical variant <ai> and the various non-standard transcriptions included in the Rest category in the original non-native matrix indicate that Spanish learners do have problems with diphthongization, even though the non-native matrix reflects fewer confusions (see Table 6.9).

Table 6.9 *Most frequent canonical (indicated in green) and non-canonical transcriptions of all 15 Dutch vowels produced by Spanish learners, as given by non-expert native Dutch listeners; target vowels in the columns, classified vowels in the rows, non-canonical transcriptions with deviations > 2.5% related to vowel height (in pink), vowel length (in turquoise), rounding (in yellow) and diphthongization (in teal) are also indicated; Vow = target vowel.*

Vow	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	Total
ie	68.0	0.8	0.0	44.2	0.6	0.0	0.0	0.0	0.0	1.2	0.8	0.0	0.0	0.1	0.1	7.7
uu	0.4	33.8	2.0	1.0	10.2	0.0	0.0	0.0	0.0	0.1	0.5	0.1	0.0	0.7	0.5	3.3
oe	0.1	30.3	72.4	0.1	37.0	1.6	0.3	0.0	0.0	0.0	0.3	5.0	0.0	2.0	4.0	10.2
i	13.7	1.6	0.1	45.0	3.8	0.0	5.6	0.0	0.0	3.1	1.7	0.0	0.3	0.3	0.1	5.0
u	0.5	4.0	2.3	3.9	39.6	3.0	1.7	0.1	0.1	0.0	3.6	0.1	0.0	0.1	2.9	4.1
o	0.0	0.0	4.2	0.0	2.3	81.8	0.0	0.6	0.0	0.0	0.5	20.6	0.0	0.1	1.4	7.4
e	2.1	0.0	0.1	1.6	2.0	0.0	87.3	1.9	0.8	17.4	5.1	0.0	3.7	1.0	0.3	8.2
a	0.1	0.0	0.4	0.0	1.3	5.6	1.2	64.8	14.2	0.3	0.0	2.0	0.0	1.0	2.0	6.2
aa	0.0	0.0	0.0	0.0	0.0	0.0	0.0	30.8	82.9	0.1	0.1	1.3	0.1	0.4	1.0	7.8
ee	12.3	0.3	0.0	4.1	0.0	0.0	1.4	0.0	0.0	58.5	1.1	0.0	6.3	0.1	0.0	5.6
eu	0.0	12.0	0.1	0.0	1.3	0.2	0.3	0.0	0.1	0.3	66.5	0.0	0.0	4.5	3.7	6.0
oo	0.0	16.2	16.2	0.0	1.1	7.3	0.0	0.0	0.0	0.0	9.6	68.9	0.0	4.8	19.0	9.5
ij	2.1	0.3	0.0	0.0	0.1	0.0	2.1	0.3	1.3	18.6	3.9	0.0	85.8	1.3	0.0	7.7
ui	0.5	0.5	0.4	0.0	0.2	0.2	0.3	1.3	0.4	0.3	4.6	0.0	3.5	68.8	8.2	5.9
ou	0.0	0.3	1.7	0.0	0.5	0.2	0.0	0.1	0.1	0.0	1.6	1.8	0.2	14.6	56.8	5.2

6.3.2.2 Native data

Table 6.10 shows how the Dutch vowels produced by two native Dutch speakers were transcribed by the native listeners (see Section 6.2.5). The columns present the 15 target vowels, while the rows show the transcribed vowels, reflecting overall percentages of canonical (indicated in green) and non-canonical transcriptions.

Our outcomes show that 81.4% of all transcriptions are canonical, whereas 18.6% are non-canonical. Such a low canonical percentage for the native data was unexpected, particularly because the native data from these two speakers was included in earlier studies (cf. Van der Harst, 2011; Van der Harst et al., 2014) and no anomalies were reported in their speech in comparison to the speech of the rest of the speakers in the native database used in those studies.

Table 6.10 shows that the target vowel with the highest percentage of canonical transcriptions is <oe> (98.6%), whereas the target vowel with the lowest percentage of canonical transcriptions is <e> (53.7%). The majority of the non-canonical transcriptions seems to be related to the non-native confusion patterns also found in the non-native matrix of the listener data (see Table 6.9). For example, the target vowel <i> is transcribed as <ie> (17.1%) (confusion related to vowel height), the target <ee> as <e> (6.0%) (vowel length) and the target <uu> as <ui> (3.8%) (diphthongization). Confusion patterns associated with rounding seem to be distributed over several vowels. Front unrounded vowels were perceived as front rounded vowels: <e> as <u> (44.4%), <i> as <uu> (10.5%), and <ij> as <ui> (21.10%). Rounding is also involved in perceiving front rounded vowels as back rounded vowels, i.e. <uu> as <oe> (23.1%) and in perceiving <a> as <o> (22.2%). The latter distinction also involves height. In sum, it seems that problematic features in perceiving non-native vowels recur in perceiving native vowels, when the latter are mixed with non-native data.

Table 6.10 *Most frequent canonical (indicated in green) and non-canonical transcriptions of all 15 Dutch vowels produced by native Dutch speakers, as given by non-expert native Dutch listeners; target vowels in the columns, classified vowels in the rows, non-canonical transcriptions with deviations > 2.5% related to vowel height (in pink), vowel length (in turquoise), rounding (in yellow) and diphthongization (in teal) are also indicated; Vow = target vowel.*

Vow	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	Total
ie	92.5	0.0	1.4	17.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.9
uu	3.7	65.4	0.0	10.5	15.6	0.0	0.0	0.0	0.0	3.0	1.5	0.0	0.0	0.0	0.0	5.4
oe	1.2	23.1	98.6	2.6	2.6	19.8	0.0	0.0	0.0	1.5	1.5	7.0	0.0	0.0	0.0	10.5
i	2.5	0.0	0.0	57.9	0.0	0.9	1.8	0.0	0.0	4.5	0.0	0.0	0.0	0.0	0.0	4.6
u	0.0	5.8	0.0	6.6	80.5	1.9	44.4	0.0	1.22	3.0	1.5	0.0	0.0	0.0	0.0	9.0
o	0.0	0.0	0.0	0.0	0.0	77.4	0.0	22.2	0.0	0.0	0.0	2.8	0.0	0.0	0.0	9.0
e	0.0	0.0	0.0	0.0	0.0	0.0	53.7	0.0	0.0	6.0	0.0	0.0	1.1	0.0	0.0	3.1
a	0.0	0.0	0.0	0.0	0.0	0.0	0.00	77.8	1.22	0.0	0.0	0.0	0.0	1.1	1.7	5.3
aa	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	96.3	0.0	0.0	0.0	0.0	0.0	0.0	7.1
ee	0.0	0.0	0.0	2.6	0.0	0.0	0.0	0.0	0.0	75.8	0.0	0.0	1.1	0.0	0.0	4.8
eu	0.0	1.9	0.0	1.3	1.3	0.0	0.0	0.0	0.0	4.5	94.0	0.0	0.0	0.0	0.0	6.2
oo	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5	90.1	0.0	1.1	10.5	6.5
ij	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5	0.0	0.0	76.7	0.0	0.0	6.3
ui	0.0	3.8	0.0	1.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	21.1	96.7	0.0	9.9
ou	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.22	0.0	0.0	0.0	0.0	1.1	87.7	4.2

6.3.3 Comparison acoustic and listener data

In this subsection we focus on the non-native data and compare the results of the acoustic data presented in the current study with the outcomes of an earlier perception study (cf. Burgos et al., submitted b). We first investigate the outcomes of the Spanish learners as a group and then examine individual differences across learners.

6.3.3.1 Spanish learners as a group

In Table 6.11, we compare the percentages of canonical classifications per target vowel of the acoustic data in the classification condition “Group” using

F1, F2 and duration (see Table 6.6), with the percentages of the native listeners' canonical transcriptions per target vowel (see Table 6.9).

A paired-samples t-test showed a significant difference between the acoustic data and the listener data ($t(14) = 2.31, p = .037$). The average canonical percentages for the acoustic data ($M = 74.7, SD = 13.3$) were higher than for the listener data ($M = 65.4, SD = 16.4$).

Remarkably, for five Dutch target vowels, namely <ie>, <oe>, <o>, <e> and <aa> (counterparts of the five Spanish core vowels /i, u, o, e, a/ respectively), the difference between the percentages of canonical classifications/transcriptions for the acoustic and listener data was fairly small (see Table 6.11), which means that the statistical classifier and the native listeners coincide to a large extent in classifying/perceiving the learner vowel realizations of these target vowels as canonical.

At the same time, there is a considerable difference ($\Delta > \pm 25.0\%$; see Table 6.11) between the acoustic and listener data for four target vowels, namely, for the short monophthongs <i> and <u>, for the long mid vowel <ee>, and for the diphthong <ui>. For the target vowels <i>, <u> and <ee>, the percentages of canonical classifications for the acoustic data are much higher than those for the perception data, indicating that many of the learners' vowel realizations of these three vowels were automatically classified as canonical on the basis of their acoustic measurements. These discrepancies show that the statistical classifier was able to classify the learner realizations of these vowels on the basis of acoustic properties that native listeners were not able to decode.

The opposite applies to the target front rounded vowel <ui> (see Table 6.11), which received a much higher percentage of canonical transcriptions for the listener data (68.8%) than for the acoustic data (41.8%). This indicates

that human vowel recognition was more accurate than the statistical classifier in perceiving the target vowel <ui>.

Table 6.11 *Percentages of canonical classifications/transcriptions per target vowel and most frequent vowel confusions, resulting from the acoustic and the perception study, confusions related to vowel height (H, in pink), vowel length (L, in turquoise), rounding (R, in yellow) and diphthongization (D, in teal) are also indicated; Vowel = target vowel, AS = acoustic study, PS = perception study, Δ = difference between the percentages of canonical classifications/transcriptions of both studies, Conf= most frequent vowel confusions per study, %Can = average percentage of canonical classifications/transcriptions.*

Vowel	ie	uu	oe	i	u	o	e	a	aa	ee	eu	oo	ij	ui	ou	%Can
AS	67.9	53.6	72.7	76.8	71.4	83.9	90.0	83.9	83.9	85.5	78.2	83.9	82.1	41.8	63.6	74.7
PS	68.0	33.8	72.4	45.0	39.6	81.8	87.3	64.8	82.9	58.5	66.5	68.9	85.8	68.8	56.8	65.4
Δ	-0.1	19.8	0.3	31.8	31.8	2.7	2.7	19.1	1.0	27.0	11.7	15.0	-3.7	-27.0	6.8	9.3
Conf-AS	i	oe	u	ie	oe	oo	ee	aa	a	ie	ui	ou	ee	ou	ui	
	21.4	21.4	9.1	14.3	14.3	8.9	3.6	7.1	14.3	9.1	16.4	7.1	8.9	21.8	12.7	
	H	R	R	H	R	L	L	L	L	H	D	D	H	R	R	
Conf-PS	i	oe	oo	ie	oe	oo	I	aa	a	ij	oo	o	ee	ou	oo	
	13.7	30.3	16.2	44.2	37.0	7.3	5.6	30.8	14.2	18.6	9.6	20.6	6.3	14.6	19.0	
	H	R	L	H	R	L	H	L	L	D	R	L	H	R	H	

Table 6.11 also shows the most frequent vowel confusions per target vowel and the features which were (erroneously) employed by the Spanish learners, namely, vowel height, vowel length, rounding and diphthongization. The rows Conf-AS and Conf-PS in Table 6.11 indicate that nine of the 15 target vowels produce the same frequent vowel confusions in the acoustic and listener data (i.e., the target vowels <ie>, <uu>, <i>, <u>, <o>, <a>, <aa>, <ij> and <ui>), whereas this is not the case for six target vowels, namely, <oe>, <e>, <ee>, <eu>, <oo> and <ou>. The target back rounded vowel <oe> was frequently classified as the front rounded vowel <u> in the acoustic data, whereas it was

perceived as the back rounded long mid vowel <oo> by the native listeners, which means that the discrepancy between the acoustic and the listener data relies on a difference in weighing front rounding and vowel length. The target front unrounded vowel <e> was often classified as the front unrounded long mid vowel <ee> in the acoustic data, but classified as the front unrounded <i> by the listeners, which indicates a difference in weight assigned to vowel length and vowel height. A similar interpretation applies to the target long mid vowel <ee>, which was frequently classified as the high front vowel <ie> in the acoustic data, but as the front unrounded diphthong <ij> in the listener data. This indicates that the difference here relates to vowel height and diphthongization.

The front rounded long mid vowel <eu> was usually classified as the front rounded diphthong <ui> in the acoustic data, but often perceived by the listeners as the back rounded long mid vowel <oo>, reflecting a disparity related to diphthongization and rounding. Similarly, the back rounded long mid vowel <oo> was frequently classified as the back rounded diphthong <ou> in the acoustic data, but perceived as the back rounded monophthong <o> by the listeners, indicating that the disparity here relates to diphthongization and vowel length. Finally, the target back rounded diphthong <ou> was frequently classified as the front rounded diphthong <ui> in the acoustic data, but as the long mid vowel <oo> by the listeners, suggesting, different weightings for the features of rounding and vowel height. In sum, larger discrepancies seem to be caused by differently weighing the competing features involved. While the four distinctive features (see Table 6.1) are clearly involved, their individual impact varies, as exemplified in Table 6.11.

6.3.3.2 *Individual differences across learners*

This subsection discusses the individual patterns found in the acoustic and listener data obtained from the classifications/transcriptions of the Dutch vowels produced by 28 Spanish learners, and the way their individual performance is related to their background characteristics, including their CEFR level, length of residence and daily use of Dutch. We first consider individual differences across learners in the acoustic data and in the listener data separately. Subsequently, we compare the outcomes of the acoustic data and the listener data for each individual learner.

6.3.3.2.1 *Acoustic data*

The dissimilarities among the Spanish learners were computed by using a matrix of 15 columns by 15 rows, giving a vector of 225 cells per learner. The analysis resulted in consistent clustering into four groups, irrespective of the clustering method used. We applied the R package *pvc* (Suzuki & Shimodaira, 2006) for a hierarchical cluster analysis with multiscale bootstrapping ($n = 1000$), using Euclidean distances and Ward's method. Two types of probability values are available: approximately unbiased (AU) p -value and bootstrap probability (BP) value, AU is a better approximation. High p -values indicate strong, certain clusters. The values vary between 0 and 100. The result of the hierarchical cluster analysis for the acoustic data is shown in Figure 6.2.

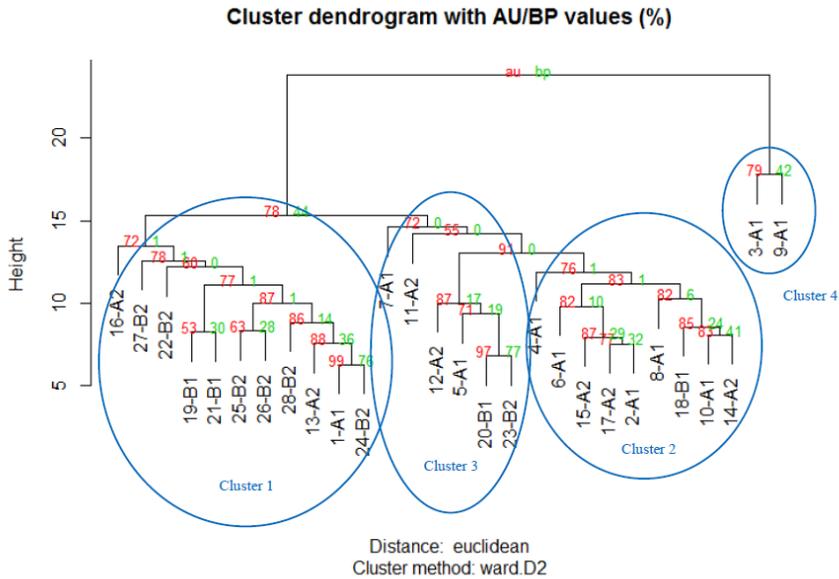


Figure 6.2 Cluster analysis of the 28 Spanish learners (including their corresponding CEFR language proficiency level) based on the percentages of canonical and non-canonical classifications per target vowel obtained from the acoustic data.

The AU values (in red) in Figure 6.2 show that three of the four clusters are not entirely separate, whereas the fourth cluster is clearly separated from the rest. There are similarities between three clusters (clusters 1, 2 and 3) and therefore between individual learners. But what are the differences between clusters? Figure 6.2 shows that the main division is between the three lower clusters (clusters 1, 2 and 3) and the fourth, higher, cluster (cluster 4). Also, a subdivision can be noted between cluster 1 and clusters 2 and 3, and between cluster 3 and cluster 2. Does this clustering result from proficiency differences? Learners with higher proficiency are likely to show a greater consistency in the production of target L2 segments, while learners with lower proficiency will tend to show more variability in their production confusion

patterns (Cutler, 2012). Table 6.12 presents the means and ranges of the percentages of canonical classifications for the Dutch vowels produced by the Spanish learners and their CEFR levels, in each of the four clusters.

Table 6.12 *Means and standard deviations (SD) for the percentages of canonical classifications per cluster for the acoustic data, including number of Spanish learners per CEFR language proficiency level; %Can = percentages of canonical classifications.*

Cluster	Mean %Can	Range %Can	CEFR levels				
			A1	A2	B1	B2	Overall
Cluster 1	80.5 (8.1)	62.0–90.0	1	2	2	6	11
Cluster 2	73.5 (8.2)	59.0–86.0	5	3	1	0	9
Cluster 3	71.6 (5.2)	66.0–79.0	2	2	1	1	6
Cluster 4	66.6 (1.7)	65.0–68.0	2	0	0	0	2

Table 6.12 shows that the highest mean percentage of canonical classifications is observed in cluster 1. The range in canonical percentages in cluster 1 overlaps with that of clusters 2 and 3, and even with that of cluster 4. This shows that it is not only the percentages of canonical classifications that determine clustering but also the distributions and percentages of the non-canonical classifications. Most of the Spanish learners are in cluster 1. When proficiency level is considered, we see that cluster 1 contains learners at all four levels, and the highest number of learners with a B2 level. Cluster 2 contains the majority of A1 learners, but also three A2 and one B1 learner. Cluster 3 does not have as many learners as clusters 1 and 2, but it has learners at all four levels. Finally, cluster 4 contains only two learners, both at A1 level.

We next consider the problems faced by learners in each of the clusters. Cluster 1 is characterized by learners with no serious problems associated with vowel length and front rounding, and with few difficulties related to vowel

height and diphthongization. Cluster 2 contains learners with problems related to height, diphthongization and front rounding, particularly in the <uu>-<u> contrast. Some of these learners also have difficulties with length; their vowel realizations are too long. The difficulties with Dutch vowels for learners in cluster 3 are similar to those observed in cluster 2, but more salient. There is great variability in the vowel confusion patterns associated with the learners in this cluster. The great majority have problems with all four distinctive features: height, length, rounding and diphthongization. They all appear to apply Spanish-like diphthongization (i.e., combining two full vowels) when realizing the long mid vowels and diphthongs.

One of the learners in cluster 1 (learner 1) rather surprisingly has an A1 proficiency level. However, this female learner received the highest percentage of canonical classifications of all the Spanish learners. An explanation for this outcome may be found in this learner's language background. She is a Spanish/Catalan bilingual who had been living in the Netherlands for six months and used Dutch daily (eight hours approximately). She was an MA student of Translation and Interpreting Studies who also spoke English and French at B2 level, and Arabic at A2 level. The fact that she spoke French is perhaps relevant, as French has front rounded vowels. Learner 28, also in cluster 1, is a male post-doctoral researcher who had been living in the Netherlands for three years. Surprisingly perhaps, his use of Dutch was rather limited (approximately four hours a day), especially if we take into account his B2 proficiency level and the fact that, of all the learners in cluster 1, he received one of the highest average percentages of canonical classifications. He was fluent in English (C2 level), French (C2 level) and German (C1 level). French and German have front rounded vowels, which may account for the phonological accuracy of this learner's Dutch vowel productions. Interestingly, for both of these exceptional learners in cluster 1 (learners 1 and

28), Dutch was their L3 or additional language (La). It is said that prior linguistic knowledge in multilinguals can be useful in the acquisition of an La (De Angelis, 2007, p. 130). Our outcomes suggest that speaking French and German, languages that have front rounding, may help learners to master front rounded vowels in an La, for example, in Dutch.

One B1 learner in cluster 2 seems to provide evidence for the suggestion that phonology acquisition does not always progress along with foreign language proficiency, (cf. Burgos, Cucchiarini, Van Hout, & Strik, 2014a). Further examination of this female learner's background (learner 18) does not offer an explanation for the low average percentages of canonical classifications she receives. She is a professional in human ecology who spoke English (C2 level) and French (C1 level). Her length of residence in the Netherlands was ten years, and her use of Dutch was low, namely four hours daily. She appeared to have problems with vowel length, vowel height and diphthongization and, perhaps most strikingly, showed an overreliance on front rounding which led to numerous vowel confusions. Her prior knowledge of French did not seem to help her produce native-like Dutch vowels. Her strategy may have been to apply front rounding for most Dutch vowels, that is, also where this was not appropriate.

The learner with the highest average percentage of canonical classifications in cluster 2 is an A1 learner (learner 2). She had been living in the Netherlands for three years. Her self-estimated use of Dutch on a daily basis was approximately 14 hours. She was a translator and fluent in English (C1 level), German (C2 level) and Italian (C1 level), which may explain her phonological skill when producing Dutch vowels.

The only B2 learner in cluster 3 represents an exceptional case. Learner 23, whose length of residence was 12 years and who used Dutch on a daily basis (eight hours approximately), is a female B2 learner who had a low-

intermediate level in English (B1 level). She had severe problems with extreme diphthongization, vowel height and vowel length, and difficulties with the front rounded vowels, in particular <uu> and <u>. This learner reported to the first author that she was fired because customers could not understand her Dutch.

Cluster 4 contains only two learners. Learner 9 is a female MA student of media studies who had been living in the Netherlands for one month. She reported using Dutch daily (6 hours approximately). She spoke English (C1 level), Portuguese (B1 level) and Catalan (B1 level). She had problems with vowel height and her long mid vowels and diphthongs were extremely diphthongized. The average percentage of canonical classifications for learner 3 was lower than for learner 9. The other learner in this cluster (learner 3) is a male university employee who had been living in the Netherlands for ten years, and did not use Dutch very much (two hours a day). He was fluent in English (C2 level) and German (C1 level). He had difficulties associated with diphthongization, vowel height and front rounding. His knowledge of German, which contains front rounded vowels, did not seem to help when producing the Dutch front rounded vowels, as attested by an evident <uu>-<u>-<oe> confusion.

Prior linguistic knowledge of other languages, especially languages with front rounding, seem to contribute to being able to produce Dutch vowels (more) accurately. In this respect, it should be noted that the B2 learner in cluster 3 (learner 23) did not speak any other foreign language (than Dutch) which has front rounded vowels.

6.3.3.2.2 Listener data

We computed dissimilarities among the speakers by using the original matrix of 15 columns by 20 rows (cf. Burgos et al., submitted b), giving a vector of

300 cells per speaker. A consistent clustering in three groups was found, regardless of the clustering method used. To exclude the noisy impact of the many cells with rather low frequencies, we excluded those cells in the matrix whose average across the informants was less than 5% of the classifications. The result was a set of 42 cells, a number that obviously is higher than the 15 cells with canonical transcriptions. We again applied the R package pvcust (Suzuki & Shimodaira, 2006) with multiscale bootstrapping (n = 1000), using Euclidean distances and Ward’s method. The result of the hierarchical cluster analysis for the listener data is displayed in Figure 6.3.

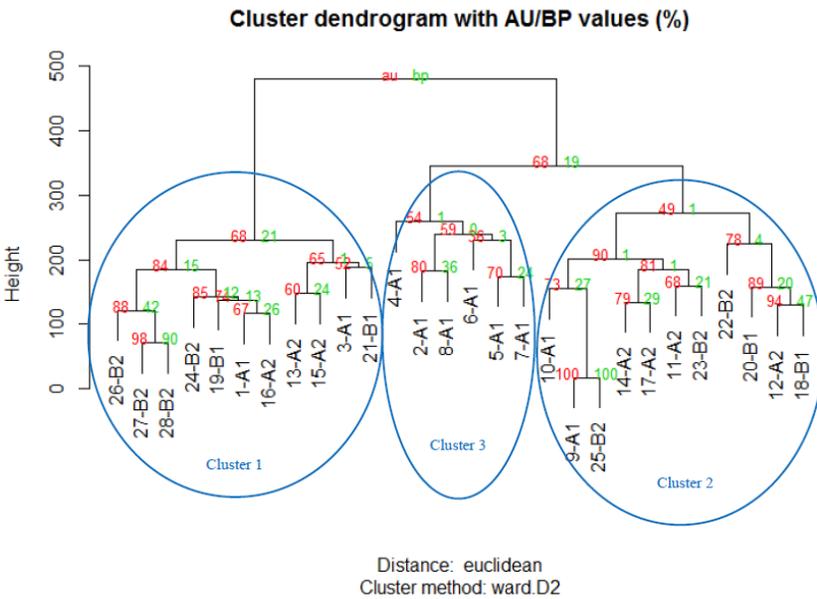


Figure 6.3 Cluster analysis of the 28 Spanish learners (including their corresponding CEFR language proficiency level) based on the percentages of canonical and non-canonical transcriptions per target vowel obtained from the listener data.

Figure 6.3 shows a different clustering than that observed in Figure 6.2. The AU values (in red) in Figure 6.3 show that the three clusters are not perfectly distinctive, pointing out that there are similarities between the clusters. But what are the differences between the three clusters? Figure 6.3 shows that the main division is between cluster 1, and clusters 2 and 3. Can this clustering be explained by proficiency differences? Again, L2 learners with higher proficiency can be presumed to have a greater consistency in the realization of phonemic target phones, resulting in higher intelligibility, whereas learners with lower proficiency will likely produce more variable input, resulting in less intelligible realizations (Cutler, 2012, p. 386). Table 6.13 presents the means and ranges of the percentages of canonical transcriptions and the CEFR language proficiency levels of the Spanish learners in each of the three clusters.

Table 6.13 *Means and standard deviations (SD) for the percentages of canonical transcriptions per cluster for the listener data, including number of Spanish learners per CEFR language proficiency level; %Can = percentages of canonical transcriptions.*

Cluster	Mean %Can	Range %Can	CEFR levels				
			A1	A2	B1	B2	Overall
Cluster 1	71.8 (9.0)	56.8–84.8	2	3	2	4	11
Cluster 2	55.5 (5.0)	47.9–64.3	2	4	2	3	11
Cluster 3	43.8 (8.1)	34.0–54.5	6	0	0	0	6

Cluster 1 is associated with the highest average percentage of canonical transcriptions, as shown in Table 6.13. The range in percentages of canonical transcriptions overlaps to some extent with the two other clusters. Clusters 2 and 3 clearly overlap in this respect. These outcomes exemplify that it is not

only the percentages of canonical transcriptions that matter, but also the percentages of non-canonical transcriptions. When it comes to proficiency, cluster 1 has the highest number of learners with a B2 level, but also two A1 learners. Cluster 2 has learners at all four levels. Cluster 3 contains the majority of A1 learners, but also one B2 learner, which suggests that L2 phonology acquisition does not always progress along with foreign language proficiency.

How are the clusters related to the transcriptions? Cluster 1 learners evoke higher percentages of canonical transcriptions, showing an overall better performance, particularly on the front rounded vowels <u> and <uu> and on the long mid vowels. This cluster is characterized by learners with no major difficulties with vowel length, front rounding and diphthongization, and probably on the verge of dealing with problems related to vowel height.

The distinction between clusters 2 and 3 is harder to define. Cluster 2 comprises learners who have difficulties with vowel height and with front rounding. Our outcomes indicate that these learners often realize the vowels <ie>, <uu>, <aa> and <oo> with longer duration. The learners' difficulties with Dutch vowels in cluster 3 are similar to those found in cluster 2, but much more salient. That is, problems with vowel height and particularly with front rounding are more severe for most learners in cluster 3. The duration of <ie>, <uu>, <oe>, <e>, <aa>, <ee> and <oo> are longer. And most importantly, all learners from cluster 3 appear to resort to extreme diphthongization when producing long mid vowels and diphthongs.

Two learners in cluster 1 have an A1 proficiency level, namely learner 1 and learner 3. The background information of these learners has already been commented on in the discussion of the cluster analysis of the acoustic data. Here again, our outcomes seem to suggest that speaking French and/or

German, languages that have front rounding, helps in mastering front rounded vowels in an L2, such as Dutch.

The three B2 learners in cluster 2 show that L2 phonology acquisition does not always reflect the level of foreign language proficiency. Learner 25 spoke English (B2 level) and her length of residence in the Netherlands was three years. Her use of Dutch was high, namely 10 hours on a daily basis. She was learning Dutch pronunciation with the help of a Dutch speech therapist at the time of the recording because she had problems being understood by native Dutch listeners. Learner 25 in cluster 2 appeared to have extreme diphthongization. She also had severe problems with front rounded vowels, particularly <u> and <uu>. A similar situation applies to learner 22 in cluster 2. She was a female B2 learner, fluent in English (C2 level), and had been living in the Netherlands for ten years. She used Dutch for an average of six hours a day. She had difficulties with the front rounded vowels, especially <u> and <uu>, and with vowel height and length.

Cluster 3 contains only A1 learners. Learner 4 is associated with the lowest average percentage of canonical transcriptions of all 28 Spanish learners, followed by learner 5, also included in cluster 3. Learner 4 is a female nurse. She had been living in the Netherlands for seven months and used Dutch on a daily basis (11 hours approximately). She was fluent in English (B2 level). She had severe problems with vowel height, vowel length, front rounding and diphthongization. Learner 5 is a male research technologist whose length of residence in the Netherlands was 7 months. He used Dutch for an average of six hours a day. He was fluent in English (B2 level). He also had severe problems with vowel height, vowel length, front rounding and diphthongization, but to a lesser extent than those observed for learner 4.

It should be noted that both B2 learners in cluster 2 and both A1 learners in cluster 3 did not speak any other foreign languages with front rounded

vowels, like French or German. This means they could not benefit from existing linguistic knowledge to help their acquisition of Dutch front rounded vowels.

We can conclude that the primary distinction among the three clusters can be related to the front rounded vowels. Our outcomes clearly show that the new feature of front rounding is affected by the L1 feature of back rounding, which leads Spanish learners to produce the front rounded vowels <u> and <uu> as the back rounded <oe>. Recurring pairwise confusions for <a>-<aa> and <i>-<ie> are detected in all three clusters, although learners in cluster 1 appear to perform considerably better when making these vowel distinctions. Although diphthongization does not cause serious difficulties and seems to compensate for problems with vowel length, extreme diphthongization, nevertheless, can lead to intelligibility problems.

6.3.3.2.3 Comparison acoustic and listener data

This subsection compares the outcomes of the acoustic and listener data per individual Spanish learner. Table 6.14 presents the mismatch between the outcomes, based on the clustering analyses, and on the average percentages of canonical classifications/transcriptions per individual learner.

Table 6.14 Mismatch between the acoustic and listener outcomes based on clustering, and on the average percentages of canonical classifications/transcriptions per Spanish learner with their corresponding CEFR language proficiency level, match (in green), mismatch of one cluster (in orange) and mismatch of two clusters (in red) are indicated; SL = Spanish learners, %Can = percentages of canonical classifications/transcriptions, AD = acoustic data, LD = listener data, Δ = difference between the average percentages of canonical classifications/transcriptions of the acoustic and listener data.

SL	CEFR	Acoustic data	Listener data	%Can AD	%Can LD	Δ	SL	CEFR	Acoustic data	Listener data	%Can AD	%Can LD	Δ
1	A1	cluster 1	cluster 1	90.0	74.1	15.9	15	A2	cluster 2	cluster 1	79.0	61.2	17.8
2	A1	cluster 2	cluster 3	86.0	54.6	31.4	16	A2	cluster 1	cluster 1	79.0	75.7	3.3
3	A1	cluster 4	cluster 1	65.0	56.8	8.2	17	A2	cluster 2	cluster 2	83.0	57.1	25.9
4	A1	cluster 2	cluster 3	59.0	34.0	25.0	18	B1	cluster 2	cluster 2	69.0	57.5	11.5
5	A1	cluster 3	cluster 3	66.0	34.3	31.7	19	B1	cluster 1	cluster 1	79.0	77.5	1.5
6	A1	cluster 2	cluster 3	72.0	46.2	31.7	20	B1	cluster 3	cluster 2	72.0	51.5	20.5
7	A1	cluster 3	cluster 3	69.0	45.0	24.0	21	B1	cluster 1	cluster 1	76.0	66.3	9.7
8	A1	cluster 2	cluster 3	72.0	48.6	23.4	22	B2	cluster 1	cluster 2	89.0	59.4	29.6
9	A1	cluster 4	cluster 2	68.0	47.9	20.1	23	B2	cluster 3	cluster 2	76.0	59.2	16.8
10	A1	cluster 2	cluster 2	69.0	51.7	17.3	24	B2	cluster 1	cluster 1	86.0	74.4	11.6
11	A2	cluster 3	cluster 2	68.0	58.5	9.5	25	B2	cluster 1	cluster 2	76.0	64.3	11.7
12	A2	cluster 3	cluster 2	79.0	49.2	36.8	26	B2	cluster 1	cluster 1	90.0	73.4	16.6
13	A2	cluster 1	cluster 1	62.0	62.5	0.5	27	B2	cluster 1	cluster 1	79.0	83.3	-4.3
14	A2	cluster 2	cluster 2	72.0	54.0	18.0	28	B2	cluster 1	cluster 1	79.0	84.8	-5.8

Table 6.14 shows that while there is a match between two clusters (acoustic data and listener data) for 15 learners, a mismatch of two clusters was found for only two learners. A mismatch of one cluster was observed for 11 learners. Mismatches were also found in terms of the degree of difference between the average percentages of the canonical classifications/transcriptions of the acoustic and listener data for the individual learners. Table 6.14 shows the average percentages of the canonical classifications/transcriptions for the acoustic and the listener data, and the difference between the two outcomes.

As noted earlier in relation to the comparison between the acoustic and listener data of the learners as a group (see Section 6.3.3.1), the average percentage of canonical classifications for the acoustic data (74.7%) is higher than for the listener data (65.4%). Almost all differences are positive, except for two learners, showing that the acoustic classification was more successful than the listener classification. The highest difference between the average percentage of canonical classifications for the acoustic data and for the listener data was observed for learner 12 ($\Delta = 36.8$), but there are more learners with high difference scores. The correlation between the two sets of percentages ($r(28) = .605$, p (two-tailed) = .001) is significant, but not high.

A closer examination of the individual patterns of learner 12 found in the acoustic and listener data, and of her background characteristics, including her CEFR language proficiency level, can help us to understand what the reasons are for such a striking difference. Learner 12 is a Spanish/Catalan bilingual who had been living in the Netherlands for six months and used Dutch daily (13 hours on average). She was working in the pharmaceutical industry and was fluent in English (C1 level), German (B2 level) and French (A2 level). The statistical classifier appears to have classified many of her Dutch vowel realizations as canonical. Her prior linguistic knowledge of other foreign languages might have contributed to her accurate production of Dutch vowels, and proficiency in German and French might have helped her in producing front rounded Dutch vowels accurately. Conversely, the average percentage of canonical transcriptions she received is rather low, indicating that the native Dutch listeners were not always able to decode the acoustic properties of her Dutch vowel realizations. An inspection of the canonical transcriptions for this speaker reveals that she had severe problems with rounding in some Dutch vowels, namely the front rounded vowels <uu> (0.0%), <u> (6.06%) and <eu> (38.1). Back rounding was also problematic, as attested by low canonical

percentages for <o> (40.63%) and <oo> (48.39%), although her production of <ou> (52.5%) was relatively successful. Difficulties with vowel height (e.g., the target vowel <i> (20.59%)) and with extreme diphthongization (e.g., the target vowel <ee> (9.38%)) were also evident. However, not all vowel realizations produced by learner 12 were inaccurate. Her realizations of the target vowels <e> (95.83%) and <aa> (89.66%), which are similar to the Spanish vowels /e, a/, as well as of the vowel <ui> (90.32%) were excellent. In sum, the outcomes of the listener data indicate that learner 12 shows a great variability in her production of Dutch vowels: the production of some vowels was poor, whereas other vowels were accurately produced, reaching near-native canonical percentages. She also shows great variability in the way she applies acoustic features. For instance, she applies front rounding proficiently when producing the front rounded vowel <ui>, reaching a near-native pronunciation, while she is not able to apply this feature properly when realizing the front rounded vowels <uu> and <u>.

The variability observed in the production patterns of learner 12 is not an exception. Interestingly, such variability seems to be present in the features associated with L2 vowel contrasts learners master predominantly. For instance, additional analyses showed that some learners seem to focus on the feature of vowel length first, which will help them to make the <a>-<aa> contrast (based on vowel height and duration) (e.g., learner 10, (A1 proficiency level) with 88.9% for <a> and 79.3% for <aa>), whereas others focus on the feature of rounding, which is necessary to produce the <ij>-<ui> distinction (e.g., learner 8, (A1 proficiency level) with 92.31% for <ij> and 90.63% for <ui>). As a result, we observe considerable variability *within* learners and *across* learners.

6.4 Discussion

The present study set out to compare the acoustic properties of Dutch vowels produced by adult Spanish learners and the perception of these vowel productions by a varied and extensive group of non-expert native Dutch listeners. To this end, we compared statistical vowel classifications obtained from the acoustic properties of the Dutch vowels produced by Spanish learners with human vowel recognition based on the transcriptions of the same Spanish-Dutch vowel productions by a large and varied group of native Dutch listeners.

An additional aim was to explain individual differences and variability in L2 vowel realizations across Spanish learners by investigating individual patterns at the production and perception levels. To establish these individual patterns, we examined the learners' proficiency level in Dutch, as well as factors that could play a role in L2 phonology acquisition, and particularly in L2 vowel accuracy, such as prior linguistic knowledge in multilinguals, length of residence and daily use of Dutch.

Our outcomes, presented in the non-native matrix (see Table 6.6) show high variability in the learners' vowel productions. The variety in, and high percentages of, non-canonical classifications assigned by the listeners indicate that the Dutch vowels produced by the Spanish learners were classified on the basis of their acoustic properties as different vowels than the target vowels. The highest variability in non-canonical classifications was found for the target front rounded vowel /œy/ (<ui>), which does not occur in Spanish (new vowel), whereas the lowest variability was observed for the target vowel /a:/ (<aa>) (similar to the Spanish /a/). The non-canonical classifications are related to vowel height, length, rounding and diphthongization. Conspicuous asymmetrical confusions were noted in the contrasts /i/-/i/ (<i>-<ie>) (based on vowel height) and /a/-/a:/ (<a>-<aa>) (based on vowel height and vowel

length), in which the vowels /i/ and /a:/ (similar to the Spanish /i/ and /a/ respectively) are frequently classified by the statistical classifier, more than /ɪ/ and /ɑ/. Vowel confusions related to rounding were reflected in the non-canonical classifications of the target front rounded /y/ (<uu>) and /ʏ/ (<u>), two new vowels which are frequently confused with each other, and particularly with the back rounded vowel /u/ (<oe>) (similar to the Spanish /u/). Similarly, the target front rounded new diphthong /œy/ (<ui>) is often classified as the back rounded diphthong /ɔu/ (<ou>). It should be remembered that Spanish does not have front rounding, as all rounded vowels in Spanish are back vowels (/o, u/) (Hualde, 2005). This could explain why the Spanish learners produce Dutch vowels /y/, /ʏ/ and /œy/ – which are new vowels to them – as back and rounded vowels. As to vowel confusions related to diphthongization, the target long mid vowel /ø:/ (<eu>) is often classified as the diphthong /œy/ (<ui>), showing evidence of extreme diphthongization in the learners' realizations.

According to Flege's (1995) Speech Learning Model (SLM), a new phoneme category may be hard to acquire when it seems similar to an existing L1 category. Adult L2 learners may use a single L1 category for two L2 phones classified as similar. In the context of the present study, the Dutch vowels /i, u, ɔ, ε, a:/ can be regarded as acoustically similar to the Spanish /i, u, o, e, a/ and therefore familiar to Spanish learners, whereas the remaining Dutch vowels (monophthongs: /y, ɪ, ʏ, ɑ/; long mid vowels: /e:/, /ø:/, /o:/; diphthongs: /ei, œy, ɔu/) can be considered new for Spanish learners. While the present study did not set out to test Flege's (1995) SLM, it can be concluded that some of our outcomes are in line with the model. They show that Spanish learners have problems in making the fine-grained vowel contrasts /ɪ/-/i/ and /ɑ/-/a:/ because the L2 phones in each pair are non-

contrastive in the L1, as both resemble Spanish /i/ and /a/. The long mid vowels and diphthongs are often produced differently than the monophthongs, namely, by applying Spanish-like diphthongization (i.e., combining two full vowels).

Our findings show that the statistical classifier and human vowel recognition coincide to a large extent in classifying/perceiving the learner vowel realizations of the Dutch target vowels /i, u, ɔ, ε, a:/ (see Table 6.11). This is in line with the Full Copying hypothesis suggested in Escudero's (2005) Second Language Perception Model (L2LP) (see also Van Leussen & Escudero (2015) for a revision of the L2LP model). A central assumption of the Full Copying hypothesis is that L2 learners will initially copy their L1 perception to attune L2 segments to their L1 native categories. Over time, exposure to the L2 will help L2 learners to evade their L1-learning mechanisms and to develop optimal L2 perception. We found evidence that the Dutch vowels that are best classified/perceived, namely, /i, u, ɔ, ε, a:/, are those vowels that are copies of the Spanish /i, u, o, e, a/ (see Table 6.11).

How do the statistical classifications of the learner vowel productions relate to their corresponding perceptions by native Dutch listeners? We assumed that the features of vowel height, length, rounding and diphthongization would play a pivotal role in perception also, but that their cue weightings might vary in comparison to the weightings used in production. The results supported our assumptions, as we found similarities and disparities between the production and native perception outcomes. As expected, comparable outcomes between the statistical classifier and native listeners were found for the five Dutch target vowels /i, u, ɔ, ε, a:/ (<ie>, <oe>, <o>, <e>, <aa>), because they match the five Spanish core vowels /i, u, o, e, a/ (cf. Flege, 1995) (see Table 6.11). This indicates that statistical vowel

classifications and human vowel recognition concur to a great extent. Disparities between the statistical classifier and native listeners were observed too. We found both slight and substantial differences. Slight differences were found for the target vowels /y, a, ø:, o:, εi, ɔu/ (<uu>, <a>, <eu>, <oo>, <ij>, <ou>), whereas substantial differences were seen for /I, γ, e:/ (<i>, <u>, <ee>, <ui>). These outcomes indicate that the human ear is able to process a large range of variability, as well as subtle and fine-grained characteristics of the speech signal in non-native speech.

The statistical classifier turned out to be more successful in classifying the learner realizations of the target vowels /I, γ, e:/ (<i>, <u>, <ee>) as canonical – based on their acoustic properties – than native Dutch listeners, who could not decode these properties or decoded them differently. It is important to take into account that the circumstances for the statistical classifier and native Dutch listeners were different. The statistical classifier considered all the data simultaneously, as a whole set, computing the solution with the best classification result. In contrast, the native listeners considered one stimulus at a time, at most within the context of previous stimuli, so that their classification can be considered to be more local than that of the statistical classifier. Therefore the native listeners not only had less information at their disposal, but their classifications might have been influenced by previous vowels in the set of stimuli they were presented with, which may have allowed them to adapt – and fine-tune – their perception.

Indeed, we found indications of adaptive mechanisms at work both in the statistical vowel classifications of the acoustic data and in native listener vowel recognition, depending on the vowel sets involved. Patterns of vowel confusions and problematic features found in the statistical vowel classifications of the non-native vowels in the classification condition “Total” (in which non-native and native data were pooled) recur in the classifications

of the native vowels (see Table 6.8). For example, problems related to the feature of vowel height appear in the non-canonical classifications of the target vowels /ɪ/ (<i>) and /y/ (<uu>) classified as /i/ (<ie>) and /ʏ/ (<u>) respectively, whereas difficulties related to rounding and diphthongization are evident from the non-canonical classifications /ɔ/ (<o>) and /œy/ (<ui>), corresponding to the target vowels /ɑ/ (<a>) and /ø:/ (<eu>) respectively. This pattern of performance in the statistical vowel classifications of the native data appears to indicate that the statistical classifier is data-sensitive and may have adapted or shifted its category boundaries to the ambiguous sounds of the non-native speech samples. This adaptive mechanism in boundary shift could help to understand why native front unrounded vowels (e.g., /ɑ/ (<a>)) were classified as back rounded vowels (e.g., /ɔ/ (<o>)). In addition, the improvement observed across the classification conditions “Total” (i.e., non-native and native data pooled together), “Group” (non-native and native data treated as two independent groups) and “Individual” (individual non-native data mixed with the native data group) indicates that the acoustic data set to be analyzed can alter the outcomes in an individual classification condition. Our outcomes for the three classification conditions seem to suggest that the statistical classifier is context-sensitive as it adapts to the nature of the data (non-native and/or native data) inputted to the system. The input of large amounts of non-native data with a large variability in vowel errors seems to lead to boundary shifts as the statistical classifier has to accommodate error-infused data (non-native data) which differ substantially from the “clean” data consisting of target categories only (native data) (cf. Berck, 2017).

Similar adaptive mechanisms in boundary shifts were observed for human vowel recognition. When listening to foreign-accented speech, native listeners seem to attend to phonetic details resulting from transfer from the learners’ L1, to navigate specific types of deviations in the speech signal. Recognizing

words with segmental deviations implies that listeners have to cope both with sounds that are distorted versions of the native norms, as well as with sounds that can be mapped onto distinct phoneme categories. Native listeners are required to shift their common boundaries to accommodate ambiguous non-native realizations which differ from their experience with native phoneme categories (cf. Bent et al., 2016; Cutler, 2012). In sum, adaptive mechanisms in boundary shifts were observed in both the statistical vowel classifications and in human vowel recognition.

The very high canonical vowel classifications obtained in the individual condition in the statistical multinomial regression analysis provides evidence that the vowels of the individual learners have acoustic distinctions, meaning that most vowels are not mergers. Not all learners make the same distinctions and not all distinctions are made with the same degree of distinctiveness. Our outcomes show that the variability in acquiring L2 phones is intricate. There is a great variability both *within* and *across* learners in their production of Dutch vowels, which leads to distinct patterns of vowel confusions per target vowel (cf. Bent et al., 2016; Mayr & Escudero, 2010). More specifically, there is great variability *within* learners both in their segmental deviations (cf. Wade, Jongman, & Sereno, 2007) and in the way different features (vowel height, vowel length, rounding and diphthongization) are used. Similarly, there is a wide range of variability *across* learners in their abilities and strategies to successfully produce the Dutch target vowels.

Our findings on individual differences across the 28 adult Spanish learners, both for the acoustic and listener data, seem to indicate that phonology acquisition does not always progress along with foreign language proficiency (see Figure 6.2, Figure 6.3 and Table 6.14) (cf. Burgos et al., 2014a). We have provided evidence that higher proficiency levels in Dutch (i.e., CEFR B2 level) do not guarantee success in achieving a native-like pronunciation in

Dutch. Other factors that are related to foreign language proficiency are length of residence and substantial L2 use. Earlier studies have shown that these factors do not appear to have a strong effect on L2 pronunciation accuracy (cf. Flege, Frieda, & Nozawa, 1997; Munro, 1993; Yeni-Komshian, Flege, & Liu, 2000). Of course, it is possible that additional factors such as intrinsic individual differences (e.g., mimicry ability, learning strategies), or socio-psychological factors (e.g., motivation to sound native-like, attitudes toward the target language and culture) may have played a role in the individual differences in L2 pronunciation accuracy across the Spanish learners (cf. Moyer, 2013 for a review of relevant factors in L2 phonology acquisition).

6.5 Conclusions

The aim of this article was to compare the acoustic properties of Dutch vowels produced by adult Spanish learners and the perception of these vowel productions by non-expert native Dutch listeners. We predicted that the features of vowel height, length, rounding and diphthongization would play a crucial role in native perception, but that their cue weightings might vary in comparison to the weightings used in production. The results supported our prediction, as we found similarities and disparities between the production and native perception outcomes. As expected, similar outcomes between the statistical classifier and native listeners were found for the five Dutch target vowels /i, u, ɔ, ε, a/, because they match the five Spanish core vowels /i, u, o, e, a/ (cf. Flege, 1995). This indicates that statistical vowel classifications and human vowel recognition concur to a great extent. Disparities between the statistical classifier and native listeners were observed too. We found both slight and substantial differences. Slight differences were found for the target vowels /y, ɑ, ø:, o:, εi, ɔu/, whereas substantial differences were seen for /l, ɣ, e:, œy/. These outcomes indicate that the native human ear is able to process

a large range of variability, as well as subtle and fine-grained characteristics of the speech signal in non-native speech.

An additional finding is that statistical vowel classifications and human vowel recognition processes are context-sensitive: in both contexts, classification processes are adapted to the nature of the data (i.e., non-native and/or native data) involved. Including non-native data (with a large variability in vowel realizations) in the analysis of native data led to different outcomes, suggesting that, with changes in the variability of the vowel stimuli, adaptive mechanisms in boundary shifts come into play in both statistical vowel classifications and human vowel recognition.

Our results on individual differences across the 28 adult Spanish learners, both for the acoustic and listener data, corroborate previous findings by showing that phonology acquisition does not always progress along with foreign language proficiency.

Finally, our findings indicate that variability in L2 phonology acquisition is extremely complex. It occurs at different levels: *within* and *across* learners with respect to segmental deviations per target vowel, and *within* and *across* learners with respect to the features (vowel height, length, rounding and diphthongization) they apply to produce Dutch vowels accurately.

7 Conclusion and discussion

7.1 Introduction

Joseph Conrad (1857-1924) was a Polish-born author who despite his brilliant command of English (lexicon, morphology and syntax), as evident from his literary masterpieces¹, was not able to reach a near-native level of acquisition in English pronunciation. In fact, it is said that Conrad's speech remained to some degree unintelligible to native English listeners throughout his life (cf. Celce-Murcia, Brinton, & Goodwing, 1996). The “Joseph Conrad phenomenon” exemplifies that L2 phonology acquisition is not only hard, but does not always run parallel to levels of mastery in other areas of the L2². What we found in the previous chapters seems to confirm this notion.

This chapter starts by summarizing the findings presented in the previous research chapters to answer the research questions formulated in Chapter 1 (Section 7.2). The most relevant issues to emerge from the research findings, namely, variability in individual learning paths (Section 7.3), adaptive mechanisms (both in statistical classification and human recognition) (Section 7.4), and L2 language proficiency and pronunciation (Section 7.5) are examined next. These issues and the research findings are then considered in the context of the speech perception models we introduced in Chapter 1 (Section 7.6). The limitations of this investigation and future prospects are

¹ Some of Conrad's great novels are *The Nigger of the 'Narcissus'* (1897), *Heart of Darkness* (1899), *Lord Jim* (1900), *Nostramo* (1904), *The Secret Agent* (1907) and *Under Western Eyes* (1911). Retrieved from https://en.wikipedia.org/wiki/Joseph_Conrad (date last viewed 31/08/17).

² The “Joseph Conrad phenomenon” (cf. Scovel, 1988) is in agreement with Flege's (2005b) “Doom” (no plasticity) hypothesis which holds that late/adult learners are unable to acquire the phonology of a second language in a native-like manner.

discussed in Sections 7.7 and 7.8. The closing section addresses the societal relevance of my research (Section 7.9).

7.2 Answering the research questions

The studies presented in this dissertation sought to gain insight into the pronunciation problems of adult Spanish learners of Dutch. The main aim of this investigation was to identify the most frequent segmental errors in Spanish learners' productions and their sources, and to determine how Spanish learners' productions are perceived by native Dutch listeners, in terms of intelligibility. To this end, four main research questions were formulated in Chapter 1 of this dissertation. Below, we will provide answers to the four research questions, based on the research findings presented in the previous chapters.

RQ1: What are the most frequent segmental pronunciation problems of adult Spanish learners of Dutch, and what are the sources for these pronunciation problems?

The most frequent pronunciation problems of adult Spanish learners of Dutch are related to Dutch vowels. The sources of these pronunciation problems are contrasts in vowel height (e.g., /I/-/i/), vowel length (e.g., /ɑ/-/a:/ and /ɔ/-/o:/) and front rounding (e.g., /y/-/u/ and /Y/-/u/), mainly due to native language interference. As to Dutch consonants, problems occur in single phonemes in word-initial position (e.g., /h/, /v/ and /j/) and word-final position (e.g., /t/), but predominantly in onset clusters (e.g., /sp/, /st/, /sx/ and /sl/) and coda clusters (e.g., /st/, /ft/, /tst/ and /skt/). Again, these mispronunciations are the result of interference from the native sound inventory and sometimes from English (e.g., as in the case of the English phoneme /dʒ/), and of phonotactic

constraints that apply in the native language (i.e., open syllable structure (CV) in Spanish vs. closed structure (CVC) in Dutch; differences in weighing the sonority hierarchy in defining permissible consonant clusters).

Additional findings in Chapter 2 showed that vowel errors are not only more frequent, but also more persistent and variable than consonant mispronunciations. Also, onset and coda clusters are responsible for a considerable number of insertions, substitutions and deletions. And finally, mispronunciations in both vowels and consonants are strengthened in several cases by the influence of orthography.

RQ₂: Do the Dutch vowels produced by adult Spanish learners match those of native Dutch speakers?

Dutch vowels produced by adult Spanish learners do not match, in terms of either duration or spectral values, those produced by native Dutch speakers. This is due to pervasive L1 constraints. The average durations of learner realizations are consistently longer than those of the native realizations. Nevertheless, learners make a distinction between short and long vowels that neatly resembles the native Dutch distinction, with the exception of the short vowel /y/, which is consistently produced with a duration typical of long vowels. With respect to vowel spectral values, learners fail to produce the subtle spectral differences required to distinguish Dutch vowel contrasts based on vowel height (e.g., /I/-/i/ and /ɤ/-/y/). As a result, learners predominantly resort to duration to realize these contrasts in a non-native manner, as in the case of the native contrasts /ɑ/-/a:/, /I/-/i/ and /ɤ/-/y/.

Supplementary findings in Chapter 3 showed that the Spanish learners' Dutch vowels which least match native realizations are the nine monophthongs, in particular where vowel contrasts based on spectral features

are concerned (e.g., /ɔ/-/u/, /ʏ/-/y/, /ɛ/-/I/, /I/-/i/ and /ɑ/-/a:/. Conversely, adult Spanish learners generally do succeed in employing diphthongization when realizing the Dutch long mid vowels and diphthongs, although their vowel productions exhibit a clear Spanish-like diphthongization pattern (i.e., combining two full vowels). A particularly relevant outcome of our investigation is that a new L2 vowel category (/[\u028c\u028a]/) was established, encompassing two L2 front rounded vowels, /ʏ/ and /y/. Overall, our findings suggest the interaction of two mechanisms in adult L2 vowel acquisition processes: the formation of new vowel categories and equivalence classification, as proposed in Flege's (1995) Speech Learning Model.

RQ₃: *Are the Dutch vowels as produced by Spanish learners of Dutch intelligible for non-expert native Dutch listeners?*

The Dutch vowels produced by adult Spanish learners of Dutch are not entirely intelligible for a diverse and large group of non-expert native Dutch listeners. Dutch vowels pronounced by Spanish learners were frequently transcribed differently from their canonical forms. The numerous confusions (e.g., /ɑ/-/a:/, /I/-/i/, /y/-/u/ and /ʏ/-/u/), as evident from the high percentage of non-canonical transcriptions, point to low intelligibility of separate Dutch words spoken by adult Spanish learners, which could hamper interaction between Spanish learners and native Dutch listeners.

An additional goal of the studies reported in Chapter 4 (crowdsource study) and Chapter 5 (study using snowball sampling) was to consolidate the usefulness of the *auris populi* methodology, i.e., the method of collecting transcriptions of L2 speech by selecting a diverse and large group of non-expert native listeners. The consistency in the confusion patterns found in the two studies provides relevant information on the way the variability in Spanish

learners' pronunciation is perceived by native Dutch listeners. This consistency demonstrates that the concept of the *auris populi* constitutes a promising approach to analyze the intelligibility of L2 speech and to pinpoint problematic areas of pronunciation. The listeners' transcriptions confirmed the vowel problems and the “attractor” effect or “similarity attraction” phenomenon found in our studies based on expert annotations (Burgos, Cucchiarini, Van Hout, & Strik, 2013, 2014a; see Chapter 2) and acoustic measurements (Burgos, Jani, Cucchiarini, Van Hout, & Strik, 2014b, submitted a; see Chapter 3).

The findings in Chapters 4 and 5 confirm that native listeners can perceive deviations in L2 accented vowel realizations, even deviations that native speakers would not produce (Magen, 1998). The variability inherent in L2 learner speech implies that native listeners have to rapidly adapt their perception to different pronunciations across learners (cf. Clarke & Garret, 2004). These adaptation processes require native listeners to adjust their phoneme boundaries, to accommodate variability in non-native speech (cf. Clarke & Luce, 2005) and to enhance language processing (cf. Bradlow & Bent, 2008; Cutler, 2012). Indeed, the outcomes in Chapter 5 in particular indicate that native listeners perceptually retune their phoneme boundaries, and that this can even lead to them categorizing native vowels non-canonically.

RQ4: Do the acoustic properties of the Dutch vowels spoken by adult Spanish learners of Dutch match the perceptual assessments by natives of these learner vowel productions?

The acoustic properties of the Dutch vowels spoken by adult Spanish learners of Dutch do not completely match the native perceptual assessments of the

same learner vowel productions by a diverse group and large of non-expert native Dutch listeners. Similarities between the production and native perception outcomes were found for the five Dutch target vowels /i, u, ə, ε, a:/, because they match the five Spanish core vowels /i, u, o, e, a/ (cf. Flege, 1995). This indicates that statistical vowel classifications and human vowel recognition concur to a great extent. Disparities – both slight and substantial – were observed too. We found slight differences for the target vowels /y, α, ø:, o:, εi, ou/, whereas substantial differences were seen for /l, γ, e:, œy/. These outcomes indicate that the native human ear is able to process a large range of variability, as well as subtle and fine-grained characteristics of the speech signal in non-native speech.

An additional finding in Chapter 6 was that statistical vowel classifications and human vowel recognition processes are context-sensitive. Classification processes are adapted to the nature of the data (i.e., non-native and/or native data) involved. Including non-native data (with a large variability in vowel realizations) in the analysis of native data led to different outcomes, suggesting that, with changes in the variability of the vowel stimuli, adaptive mechanisms in boundary shifts come into play in both statistical vowel classifications (cf. Berck (2017) on this phenomenon in another data domain) and human vowel recognition (cf. Bent, Baese-Berk, Borrie, & McKee, 2016; Cutler, 2012).

In answering the four research questions in the different chapters, we investigated individual variation by examining the learners' proficiency level in Dutch, as well as factors that could play a role in L2 phonology acquisition, and particularly in L2 vowel accuracy, such as prior linguistic knowledge in multilinguals, length of residence and daily use of Dutch. Our results on individual differences across the 28 adult Spanish learners, both for the

acoustic and listener data, corroborate previous findings (cf. Burgos et al., 2014a; Flege, Frieda, & Nozawa, 1997; Munro, 1993; Yeni-Komshian, Flege, & Liu, 2000) by showing that phonology acquisition does not always progress at the same rate as foreign language proficiency in general. As to the variability in L2 vowel realizations across the individual learners, our outcomes indicate that variability in L2 phonology acquisition is extremely complex. The individual learners' performance shows that variability occurs at different levels: *within* and *across* learners in their segmental deviations per target vowel, and *within* and *across* learners in the strategies or features (vowel height, vowel length, rounding and diphthongization) they apply to produce Dutch vowels accurately.

This section has provided answers to the four main research questions. The most relevant issues to emerge from the research findings, namely, variability in individual learning paths, adaptive mechanisms (both in statistical classification and human recognition), and the relationship between L2 language proficiency and pronunciation performance, are discussed below.

7.3 Individual variability

Our findings when it comes to explaining individual differences across the learners' performance in L2 phonology acquisition have shown that all 28 adult Spanish learners of Dutch show great variability in the way they produce their Dutch vowels. The very high score of canonical vowel classifications obtained in the individual condition in the statistical multinomial regression analysis (89.1%; see Chapter 6) provides evidence that the vowels of the individual learners have acoustic distinctions, meaning that most vowels are not mergers. Learners do not all make the same distinctions and not all distinctions are made with the same degree of distinctiveness. Our outcomes

provide evidence that the variability in acquiring L2 phones is extremely intricate, and that it occurs at different levels. More specifically, there is great variability *within* learners, both in their segmental deviations (cf. Wade, Jongman, & Sereno, 2007) and in the way they use different features (vowel height, vowel length, rounding and diphthongization). This is the case for all target vowels. Also, there is a wide range of variability *across* learners in their ability to accurately produce the Dutch target vowels. The variability *within* and *across* learners leads to distinct patterns of vowel confusions per target vowel (cf. Bent et al., 2016; Mayr & Escudero, 2010)³. Learners seem to take recourse to different features such as vowel height, vowel length, rounding and diphthongization to produce Dutch vowels accurately, although their reliance on a particular feature to make a certain vowel contrast is not always appropriate. Variability is also observed in the vowel contrasts learners master first. That is, some learners with an A1 proficiency level in Dutch master the /ɑ/-/a:/ contrast first, whereas others appear to master the /ei/-/œy/ distinction at an early stage of their acquisition. In other words, there is considerable variability in the paths learners follow in their efforts to achieve an accurate pronunciation of Dutch target vowels.

Although we found great variability in individual performance, we did not investigate what the factors are that drive this variability. For decades, studies on Second Language Acquisition (SLA) have assumed that most limitations in L2 development/acquisition follow from maturational (Critical Period Hypothesis; Lenneberg, 1967), muscular (Scovel, 1988) and/or cognitive constraints⁴. However, recent studies aimed at understanding individual

³ See also Hazan, Sennema, & Faulkner (2002) who showed differences *within* and *across* Spanish learners of English in their use of audiovisual cues in the perception of sound contrasts (/b/-/v/ and /p/-/b/) which have a different phonemic status in the listeners' L1 and L2.

⁴ See Moyer (2013) for a discussion on age constraints and their effect on foreign accent.

differences in language learning have shown that cognitive, psychological and social mechanisms also play a role in the accuracy with which a second language is acquired (cf. Larsen-Freeman, 2009). The trade-off between these mechanisms may explain why, for instance, learners with the same L1 who are exposed to the same target language at the same age exhibit differences in their L2 development/acquisition. Intrinsic and extrinsic differences can account for this variation in individual performance, particularly when it comes to phonological learning (Moyer, 2013). Intrinsic individual differences relate to, for example, differences in aptitude (e.g., mimicry ability: some learners have a special talent to learn languages⁵ and to imitate accents), musical talent⁶, learning styles and strategies⁷ and gender⁸. Extrinsic individual differences may stem from socio-psychological factors, such as identity, motivation and attitudes, or from differences in experience and input. Identity is associated with (foreign) accent in that L2 learners may not want to sound native in the L2 because this may feel as rejecting their identity and

⁵ A common Dutch word to describe these talented learners is *talenknobbel*, ‘linguistic talent’.

⁶ See Gottfried (2008) and Tokuhamo-Espinosa (2003) who found that musical talent contributes to phonological learning.

⁷ *Learning style* refers to different styles of learning preferred by each individual learner (seven major learning styles can be classified into the following categories: visual (spatial), aural (auditory-musical), verbal (kinesthetic), logical (mathematical), social (interpersonal) and solitary (intrapersonal); see Kolb & Kolb, 2005). *Learning strategy* concerns goal-oriented techniques to maximize learning achievements. Learning styles and strategies are unquestionably related to personality which has been seldom investigated in empirical studies on L2 phonology (Moyer, 2013).

⁸ See Moyer (2013) for a review of studies addressing gender differences in L2 phonology research.

native culture (cf. Lybeck, 2002; Pavlenko & Lantolf, 2000)⁹. With regard to motivation, several studies have shown that highly motivated L2 learners are more likely to achieve a near-native pronunciation (cf. Bongaerts, 2005; Bongaerts, Planken, & Schils, 1995; Moyer, 1999), especially when driven by a combination of personal and professional motivation (Flege, Yeni-Komshian, & Liu, 1999; Moyer, 2007). As to attitude, we can distinguish attitude toward learning foreign languages, toward the target L2 itself, and toward its culture and community of speakers. Attitude has been examined in foreign accent research in terms of concern for pronunciation accuracy, desire to sound native, self-rating of accent, and attitudes toward the target language and culture (Moyer, 2013, p. 70; cf. Masgoret & Garner, 2003; Moyer, 1999, 2004, 2007).

Two other factors that could play a role in individual variability are experience and input. Both are often measured in terms of amount of time (weeks, months, years) of L2 exposure, and in terms of length of residence (LoR)¹⁰. However, such measures can lead to misleading assumptions, as amount of L2 exposure or LoR does not account for learners' L2 phonological accuracy, as shown in the previous chapters (especially in Chapter 6, in which the individual differences across the 28 Spanish learners were examined). As to input, it is not the *quantity* that seems to affect L2 phonology attainment, but the *quality*. Exposure and LoR as measures of amount of L2 input tell us nothing about the quality of the input received (cf. Flege, 2009, 2012). According to Flege (2012), both the quantity and quality of L2 input are

⁹ According to Pavlenko and Lantolf (2000, p. 266), L2 learners who become members of a new language (L2) and cultural community may affiliate themselves with more than one language or culture, which implies that a single L2 learner can hold multiple identities.

¹⁰ The significance of the variable length of residence (LoR) for L2 foreign accent is rather unstable across studies (see Piske et al., 2001).

essential to L2 speech learning and need to be taken into account when explaining variability in individual performance¹¹.

An illustrative example of the relevance of L2 input is provided by the case of Spanish learners of English. A frequent question this author faced when moving to the Netherlands was “Why are Spaniards so bad at speaking English, particularly compared to Dutch people?” One of the reasons can be found in the phonological differences between Spanish and English (cf. Hualde, 2005). Another fundamental reason is the kind/amount of (English L2) input received by Spaniards throughout their lives. In Spain, it is extremely hard for Spanish people to gain exposure to native English sounds when British and American programs, series and movies are always dubbed in Spanish. The same applies to learning English in Spain; learners are not familiarized with English sounds in the classroom as teachers are likely to speak Spanish in English lessons and will mostly focus on English grammar and lexicon, while neglecting listening and speaking. Furthermore, Spanish teachers’ English pronunciation tends to be poor and variable so that evident pronunciation errors are transferred to learners. As a result, Spanish learners tend to have problems speaking English and find it hard to make themselves understood to native English listeners. Previous research has underestimated the importance of L2 input (see Flege (2009) for a review), concluding that input is more important for learning the L1 than it is for learning an L2, and that variability in adult learners’ speech is due to “undefined” individual differences, rather than to the kind/amount of L2 input received (Flege, 2012). At present, the role of L2 input remains relatively understudied in SLA

¹¹ See Flege’s (2012) slides for his lecture “The role of input in second language (L2) speech learning”, presented at the VIth International Conference on Native and Non-native Accents of English, Łódź, Poland, 6–8 December 2012. Retrieved from <http://www.jimflege.com/> (date last viewed 31/08/17).

research. More research needs to be done to throw light on the way input influences individual differences in L2 learning and in establishing valid pedagogical priorities for L2 learners.

The variability inherent in L2 learners' production implies that native listeners have to adapt to different pronunciations across learners. For example, in the context of L2 vowel realizations, they have to be able to shift their category boundaries to accommodate an ambiguous vowel realization that differs from their usual expectations about phonemic categories (Cutler, 2012). These perceptual adaptation processes allow native listeners to adjust the L1 boundary between categories to accommodate variability in L2 speech, and benefit language processing (cf. Bradlow & Bent, 2008; Cutler, 2012). Adaptive mechanisms will be discussed next.

7.4 Adaptive mechanisms (statistical classification and human recognition)

The variability inherent in L2 learners' production implies that statistical classification and human recognition have to adapt to different pronunciations both *within* and *across* learners to be able to perceive intended meaning. Our outcomes provided evidence of adaptive mechanisms at work both in the statistical vowel classifications of the acoustic data and in human vowel recognition, depending on the vowel sets involved (see Chapter 6). Patterns of vowel confusions and problematic features found in the statistical vowel classifications of the non-native vowels in the classification condition "Total" (in which non-native and native data were pooled) recur in the classifications of the native vowels. For example, problems related to the feature of vowel height appear in the non-canonical classifications of the target vowels /I/ and /y/ classified as /i/ and /ɤ/ respectively, whereas difficulties related to rounding and diphthongization are evident from the non-canonical classifications /ɔ/

and /œy/, corresponding to the target vowels /ɑ/ and /ø:/ respectively (see Chapter 6). This pattern of performance in the statistical vowel classifications of the native data indicates that the statistical classifier is data-sensitive and may have adapted or shifted its category boundaries to the ambiguous sounds of the non-native speech samples. Such boundary shifts help to understand why native front unrounded vowels (e.g., /ɑ/) were classified as back rounded vowels (e.g., /ɔ/). In addition, the improvement observed across the classification conditions, from “Total” (i.e., non-native and native data pooled together) to “Group” (non-native and native data treated as two independent groups) and to “Individual” (individual non-native data mixed with the native data group) indicates that the acoustic data set under analysis can alter the outcomes in an individual classification condition (see Chapter 6). Again, the outcomes across the three classification conditions seem to suggest that the statistical classifier is context-sensitive; it adapts to the nature of the data (non-native and/or native data) that is input into the system. The input of large amounts of non-native data with high variability in vowel errors seems to lead to boundary shifts where the statistical classifier accommodates error-infused data (non-native data) which differ substantially from the “clean” data consisting of target categories only (native data) (cf. Berck (2017) who shows that machine learning algorithms are affected by infusing errors in linguistic data).

Similar adaptive mechanisms in boundary shifts were observed for human vowel recognition (see Chapters 5 and 6). Native listeners can rapidly perceive segmental deviations from the norm and can easily detect pronunciation errors by learners that native speakers are not likely to make (Magen, 1998). The variability inherent in L2 learners’ production implies that native listeners have to adapt to different pronunciations across learners. For example, when listening to foreign-accented speech, native listeners seem to attend to

phonetic details resulting from transfer from the learners' L1, to navigate specific types of deviations in the speech signal. To be able to recognize words with segmental deviations listeners have to cope both with sounds that are distorted versions of the native norms, as well as with sounds that can be mapped onto distinct phoneme categories (e.g., the Spanish learners' realizations of the target vowels /l/ and /a/ were often perceived by native Dutch listeners as /a:/ and /i/ respectively; see Chapters 5 and 6). Native listeners are required to shift their common boundaries to accommodate ambiguous non-native realizations which differ from their experience with native phoneme categories (cf. Bent et al., 2016; Cutler, 2012), and benefit language processing (cf. Bradlow & Bent, 2008; Cutler, 2012). As a consequence, native listeners may have to temporarily adapt or shift their category boundaries to the ambiguous phonemes in the non-native samples. Such processes could help to understand why native front unrounded vowels were transcribed as front rounded vowels (e.g., /ɛ/ as /y/, /l/ as /y/ and /ei/ as /œy/; see Chapter 5). Earlier work has shown that native listeners can adjust rapidly to non-native realizations (cf. Witteman et al., 2014), ignoring their long-term native representations of those realizations (Clarke & Garret, 2004). Our results seem to provide evidence of what Cutler (2012) describes as "the plasticity in adult native listeners' perception" (p. 375). Native listeners' perceptual adaptation may generate a boundary shift between categories when being exposed to non-native (cf. Clarke & Luce, 2005) and native speech, especially when vowels share a phonological feature (cf. Chládková, Podlipský, & Chionidou, 2017), such as rounding (e.g., the case of the front rounded vowels /y, y/ and the back rounded vowel /u/). Our outcomes provide evidence that the native human ear is able to process a large range of variability in the speech signal of non-native speech. In sum, adaptive

mechanisms in boundary shifts were observed in both the statistical vowel classifications and in human vowel recognition.

It should be noted that adaptive mechanisms are not only active when listening to non-native speech, but also when listening to native speakers of the same native language. Native listeners are very good at adapting rapidly to (new) speech: new words, pronunciation variability within and across talkers, dialectal variation, unfamiliar accents, and language change across time. Their perceptual adaptation relies on the plasticity of adult speech perception, through which native phonemic boundaries can be adapted as the listening situation requires in order to facilitate communication. A basic principle of native listening is thus the exceptional ability of listeners to adapt their speech processing to the great variability in native speech (Cutler, 2012).

7.5 L2 language proficiency and pronunciation

The great variability observed across learners' performance, irrespective of their language proficiency level in the L2, affects the phonological accuracy with which L2 segments are produced, and, subsequently, perceived by native listeners. Our findings on individual differences across the 28 adult Spanish learners of Dutch, both for the acoustic and listener data, seem to indicate that phonology acquisition does not always progress along with foreign language proficiency (see Chapter 6; see also Chapter 2). High proficiency level in Dutch (i.e., CEFR B2 level) does not guarantee success in achieving a native-like pronunciation in Dutch (e.g., the case of learner 23 (B2 level in Dutch) who was fired because customers could not understand her Dutch; see Chapter 6). Factors that are related to foreign language proficiency are length of residence and substantial L2 use. Earlier studies have shown that these factors do not appear to have a strong effect on L2 pronunciation accuracy (cf. Flege, Frieda, & Nozawa, 1997; Munro, 1993; Yeni-Komshian, Flege, & Liu, 2000).

Few studies have investigated the extent to which L2 language proficiency correlates with L2 pronunciation. Also, a description of the phonological requirements learners need to meet to have a certain proficiency level in the L2 is not specifically described in current standards models of L2 learning (see also Moyer (2013) for a discussion on accent within current standard models). For example, in the descriptors of the Common European Framework of Reference for Languages (CEFR)¹² phonology or phonological control receives a marginal importance as one of the six subdivisions¹³ of the subdomain linguistic competence (next to two other subdomains: sociolinguistic and pragmatic competencies), which belongs to the broad area of communicative language competence. In addition, it should be pointed out that in the CEFR descriptor for phonological control not all phonological abilities are described in the three broad proficiency levels (i.e., A (basic user), B (independent user) and C (proficient foreign language user)), as displayed in Table 7.1.

¹² See framework guidelines for the CEFR in <http://www.coe.int/en/web/common-european-framework-reference-languages/> (date last viewed 31/08/17).

¹³ All six subdivisions of the subdomain linguistic competence are general range, vocabulary range, grammatical accuracy, vocabulary control, phonological control and orthographic control. Retrieved from <https://rm.coe.int/168045b15e> (date last viewed 31/08/17).

Table 7.1 *Phonological control of the CEFR descriptors per proficiency level.*

Proficiency levels	Descriptor	
Proficient user	C2	No descriptor available.
	C1	Can vary intonation and place sentence stress correctly in order to express finer shades of meaning.
Independent user	B2	Has a clear, natural, pronunciation and intonation.
	B1	Pronunciation is intelligible even if a foreign accent is sometimes evident and occasional mispronunciations occur.
Basic user	A2	Pronunciation is generally clear enough to be understood despite a noticeable foreign accent, but conversational partners will need to ask for repetition from time to time.
	A1	Pronunciation of a very limited repertoire of learnt words and phrases can be understood with some effort by native speakers used to dealing with speakers of his/her language group.

Table 7.1 shows that no descriptor is available for a proficient user with a C2 level. Also, general comments such as “a noticeable foreign accent” (see descriptor for A2) or “occasional mispronunciations” (see descriptor for A2) lack specific information about the phonological requirements learners need to meet to have one of these six different proficiency levels. The descriptors presented in Table 7.1 recognize the communicative relevance of accent, but they mix it up with other language skills sets such as range (see descriptor for A1) or interaction (see descriptor for A2). It is obvious that phonology is not treated as an essential linguistic domain by current standard models of proficiency, such as the CEFR. The findings of this investigation have shown that L2 phonology acquisition is extremely important for the intelligibility of L2 speech and therefore for communicating successfully in an L2 environment. The findings of this investigation illustrate the complexity of individual variability as well. It is tempting to conclude that current standard models cannot present a more specific description of phonological control per

proficiency level because pronunciation competence is extremely hard to define just because of the great variability across L2 learners.

7.6 Evaluating the speech perception models

This section evaluates the findings in the context of the speech perception models that formed the framework for the current investigation. The shared tenet of L2 speech perception models such as the Speech Learning Model (SLM; Flege, 1999, 2003), the Perceptual Assimilation Model (PAM; Best, 1995) and its extension, the PAM-L2 (Best & Tyler, 2007), and the Second Language Linguistic Perception Model (L2LP; Escudero, 2005; see also Van Leussen & Escudero (2015) for a revision of the L2LP model) is that specific sound patterns of the adult learners' native phonology affect L2 speech perception (and eventually L2 speech production) (see Chapter 1, Section 1.5 for a short description of these speech perception models). Although the focus of this dissertation was not to test and compare these models, Flege's SLM was highly relevant to one of our studies (see Chapter 3) as it is the only speech model which focuses on both the perception and production of L2 speech by L2 learners. Importantly, the SLM explicitly addresses the creation of new L2 sounds, which was one of the outcomes of Chapter 3. The Full Copying hypothesis described in the L2LP model (Escudero, 2005; cf. Van Leussen & Escudero, 2015) was also used to interpret our findings. The outcomes of Chapter 3, providing confirmation of the mechanism of equivalence classification and the formation of new categories (SLM; Flege, 1995), as well as the Full Copying hypothesis (L2LP; Escudero, 2005) are presented in Table 7.2, that is the same as Table 3.2 in Chapter 3. We copied it because it nicely illustrates in detail how the vowel pronunciation patterns of the Spanish learners of Dutch can be modelled.

Table 7.2 Findings of Chapter 3 based on Flege’s Speech Learning Model and Escudero’s Second Language Linguistic Perception Model.

	L1 Spanish	Spanish-Dutch vowels	Target L2 Dutch
Equivalence classification (similar Spanish monophthongs)	/i/	/i/	/ɪ/ /i/
	/e/	/e/	/e/ /ø/
	/a/	/a/	/ɑ/ /a:/
	/o/	/o/	/ɔ/ /o:/
	/u/	/u/	/ʊ/ /u:/
New category (front round)	-	/[ɣ~y]/	/ɣ/ /y/
Equivalence classification (copy Spanish diphthongization)	/e/+i/	/e+i/	/e:/ /ø:/
	/e/+u/	/e+[ɣ~y]/	/ø:/ /o:/
	/o/+u/	/o+u/	/o:/ /ɔ:/
	/a/+i/	/a+i/	/ɛi/ /œy/
	/a/+u/	/a+[ɣ~y]/	/œy/ /ɔu/
	/a/+u/	/a+u/	/ɔu/ /o:/
Split mergers through duration	/a/	/a’-/a:/	/ɑ’-/a:/ /a’-/a:/
	/i/	/i’-/i:/	/ɪ’-/i/ /i’-/i/
	-	/[ɣ~y]’-/ [ɣ~y:]’/	/ɣ’-/y/ /ɣ’-/y/
Orthographic vowel lengthening	-	/y:/	/y/ (<uu>)
	-	/a:/	/a/ (<aa>)
	-	/e:/	/e/ (<ee>)
	-	/o:/	/o/ (<oo>)

Table 7.2 indicates how establishing new categories and making use of equivalence classifications offers adult Spanish learners – with their native 5-vowel system (/i, u, o, e, a/) (Hualde, 2005) – the possibility to realize 15 vowel distinctions (nine monophthongs: /i, y, u, ɪ, ʏ, ɔ, ɛ, ɑ, a:/; three long mid vowels: /e:/, ø:/, o:/; three diphthongs: /ɛi, œy, ɔu/) (Booij, 1995), the number of vowels Dutch has, although their realizations are not native-like.

We found that Spanish learners make more distinctions in Dutch vowels than one might expect, as evidenced by the large number, overall, of acoustic and perceptual vowel confusions we found. When the learners were added individually in the statistical vowel classification analysis, their success scores were high (89.1%; see Chapter 6). This result also shows that learners' performance is variable, as they try in diverse ways to implement the features of height, rounding, duration and diphthongization. These individual differences and the huge variability in production are not predicted by the three perception models, whose perspective is on reducing perceptual L2 categories and their variability. While these perception models are adequate predictors of the type of problems adult learners face in learning the L2 vowel system, they do not cater for the intricate interaction between the different features to be learned and the variable production within and across learners. They do not explain the extent to which individual learners produce L2 contrasts which are not sufficiently precise, subtle or fine-grained to be perceived by native listeners.

The findings in this dissertation, for both production and native perception, indicate that Spanish learners use their native five-vowel system as departing point to acquire all 15 Dutch vowels. Although we did not study orders of acquisition, by way of speculation, it seems conceivable that in acquiring all Dutch vowel categories Spanish learners would apply an order of vowel acquisition which may proceed as follows. The Spanish vowels /i, e, a, o, u/ are used to establish the Dutch vowel categories /i, ε, a:, ɔ, u/ (number of vowels = 5). A new front rounded category, namely, /[y≈y]/ is established (number of vowels = 6). Next, the Dutch long mid vowels (/e:, ø:, o:/) and diphthongs (/εi, œy, ɔu/) would be produced by connecting two existing vowel categories, including the new front rounded vowel, just like Spanish vowel combinations (number of vowels = 12). Finally, the new feature of

duration is employed to split three existing vowel categories (i.e., Spanish /a/ and /i/, and the new front rounded / $[y \approx \gamma]$ /) to distinguish new categories: three short (/a, i, $[y \approx \gamma]$ /) and three long vowels (/a:, i:, $[y \approx \gamma]:$ /) instead of the native contrasts /a/-/a:/, /I/-/i/ and /Y/-/y/ (number of vowels = 15). The most conservative estimation, excluding establishing new categories, would deliver 9 vowels (number of vowels = 5 (i.e., the Dutch vowel categories /i, ε, a:, ə, u/); number of vowels = 4 (i.e., the Dutch long mid vowels /e:, o:/ and diphthongs /ei, əu/)). It is tempting to speculate which orders of acquisition indeed occur, but to test orders of acquisition we would need to collect more elaborate data in a longitudinal study, preferably not restricted to reading a word list (see Chapter 3, Subsection 3.3.2).

7.7 Limitations

This dissertation has provided many answers, but, as is common in an investigation, there are limitations too. First, the initial aim of the project that was the starting point for this dissertation was to systematically investigate the most frequent pronunciation errors of Spanish learners of Dutch, to gain knowledge which could be used as a guideline in developing a dedicated ASR-based CAPT program¹⁴ for the Spanish L1-Dutch L2 language pair. This project eventually did not progress along the lines of a CAPT application. We would have liked to have developed a CAPT system aimed at providing sufficient practice and personalized, instantaneous feedback on pronunciation, but this is now something for the future (see also Section 7.8). Second, we note that the studies reported, particularly those focusing on the speech production of adult Spanish learners (see Chapters 2 and 3), were conducted

¹⁴ ASR stands for automatic speech recognition; CAPT stands for computer assisted pronunciation training.

with a relatively small number of participants. Including the speech data of more participants in our statistical analyses might have enhanced the representativeness of our findings. A third limitation is that using read speech only may have affected our outcomes for the speech production data, as orthography may have influenced the Spanish learners' speech production. Obtaining spontaneous speech in a semi-controlled task such as picture naming or word/sentence repetition, in which participants are not given written stimuli, would provide potentially useful additional data. A fourth and final limitation is that we did not elicit Spanish learners' productions of Spanish vowels and vowel combinations as well as their Dutch learner productions. With hindsight, this would have been useful, to determine whether the learner realizations of the target L2 phones match their own realizations of L1 sounds to a greater/lesser extent than the realizations of the target phones produced by native speakers of the L2. We suggest, therefore, that future studies in a similar vein elicit and analyze not only non-native, but also native, samples of speech from learners being investigated.

7.8 Future prospects

In the studies reported in this dissertation, the learners' speech was elicited at one moment in time. It is evident that more precise and valid conclusions about the process of acquisition can be drawn on the basis of a longitudinal study (cf. Ortega & Ibarri-Shea (2005) for an elaboration on longitudinal studies). A good example of the benefit of a longitudinal investigation is the study by Trofimovich, Lightbown, Halter, and Song (2009). In a two-year comparison of francophone Canadian learners of English L2 taking part in either an experimental comprehension-based program or a "regular" language learning program, Trofimovich et al. (2009) showed that while there were no differences between learners in the two programs after one year, their

pronunciation scores did differ at the end of the second year. Had Trofimovich et al.'s (2009) study not extended over two years, these differences would not have come to light. Although longitudinal research is prone to certain threats such as funding limitations, testing effects, and participant fallout (Munro & Derwing, 2015), it would be useful in pronunciation research in general, and certainly in future studies with similar aims to the present investigation.

Future research should further investigate individual learners' L2 phonology acquisition (cf. Bent et al., 2016; Mayr & Escudero, 2010) and, specifically, deviations in the segmentals of non-native speech and how these are perceived by native speakers (cf. Anderson-Hsieh, Johnson, & Koeler, 1992). Understanding the phonological cues native listeners use to assess sound segments as non-native segments is an essential step in establishing valid priorities and pedagogical approaches for phonological instruction in the classroom (Derwing, 2008; Derwing & Munro, 2005; Moyer, 2013).

Another future research path to follow is to develop dedicated CAPT systems. CAPT systems that make use of ASR and automatic error detection can provide relevant practice and personalized, instantaneous feedback for L2 learners who wish to improve their pronunciation any time, anywhere and in their own tempo (Ehsani & Knodt, 1998; Eskenazi, 1999, 2009; Neri, Cucchiarini, Strik, & Boves, 2002; Pennington, 1999; Witt, 1999). The "joint venture" of CAPT and ASR can help motivated learners to achieve a more native-like pronunciation in the L2. CAPT systems can clearly benefit from studies such as those reported in this dissertation.

The focus in this dissertation was on acoustic analysis and intelligibility. According to Munro and Derwing (2015, p. 15), intelligibility and comprehensibility both have a much greater impact on the effectiveness of communication than accent alone. That is, foreign-accented speech does not necessarily impede communication as long as the pronunciation of adult

learners is intelligible (Derwing & Munro, 2005). However, there are still situations in which adult learners who can communicate effectively in the L2 are judged on the basis of their foreign accent instead of their competencies. Investigating evaluative reactions to foreign-accented speech, both in informal settings and in the workplace (cf. Carlson & McHenry, 2006; Deprez-Sims & Morris, 2010; Mai & Hoffman, 2014; Moyer, 2013) is an issue that deserves further attention in research. When native speakers listen to foreign-accented speech, they do not only judge L2 speech in terms of accentedness, intelligibility and comprehensibility, but they also judge the L2 speaker who produces the speech as a social being (Moyer, 2013). Extensive research on language attitudes has shown that accents are widely associated with social values like correctness, educatedness, competence, self-confidence and intelligence (Brown, Giles, & Thakerar, 1985), with status and solidarity (Brennan & Brennan, 1981), with status and power (Cargile, 2000; Cargile & Giles, 1998) and with credibility (Lev-Ari & Keysar, 2010) (see also Moyer (2013) for a review of studies investigating evaluative reactions to non-native speech). Several investigations have shown that foreign-accented speech can evoke negative reactions in native listeners, which can be disadvantageous for successful interaction and social acceptance (Brennan & Brennan, 1981; Lippi-Green, 1997; Moyer 2013). Follow-up research on evaluative reactions to Spanish-accented Dutch speech could be done by using the paragraphs spoken by the adult Spanish learners of Dutch we investigated (see content corpus II in Chapter 1, Section 1.7).

Another promising opportunity related to L2 speech research is investigating evaluative reactions to foreign-accented speech in the workplace and whether these reactions affect communication. For example, future research could investigate the extent to which foreign-accented speech, intelligibility and language proficiency influence successful communication

and power dynamics in multinational teams (MNTs) which operate in multinational corporations (MNCs), by investigating how native speakers of Spanish communicate in Dutch or in English with both native and non-native speakers of Dutch or English as members of the same MNT. Research has shown that language can be used as a source of power in the workplace (cf. Kingston, 1996). For example, some employees who have foreign-accented speech and have not mastered the L2 at a highly proficient level (i.e., C2 level) can feel excluded, for instance, in a meeting in which co-workers (either native or non-native speakers of the L2) are able to successfully communicate in the L2 (cf. Tenzer & Pudelko, 2017). This situation often occurs in MNCs in which English as a lingua franca is used, mostly among non-native speakers of English, as the corporate ‘business’ language. According to Kankaanranta and Planken (2010), Business English as a Lingua Franca (BELF) is a linguistic resource commonly used in today’s global business environment which can be characterized as “simplified and hybridized English” (p. 392). Non-native speakers of English, depending on their native language, are likely to use BELF in variable ways in terms of word choice and sentence structure, and as a result of their individual levels of phonological attainment, for example (cf. Akkermans, Harzing, & Van Witteloostuijn, 2010). In addition, non-native speakers of English in MNCs who operate in MNTs are likely to have a (mild or heavy) foreign-accented speech (depending on their native language) when speaking BELF (Maai & Hoffman, 2014), as well as a different degree of intelligibility or language proficiency in comparison to other non-native co-workers. These linguistic differences are especially evident in headquarters-subsidary communication (cf. Harzing & Pudelko, 2014; Harzing, Köster, & Magner, 2011). It is unclear to what extent degree of foreign-accented speech, intelligibility and language proficiency in employees who are non-native speakers of English influence power dynamics

in MNTs, or co-workers and managers' reactions, attributions and behaviours. Studies that investigate such issues may be valuable for human resource managers in MNCs and can contribute to the body of language research in international business.

7.9 Societal relevance

It is hoped that the findings from this investigation can contribute to the development of specific learning tools for native speakers of Spanish who wish to improve their pronunciation accuracy in Dutch.

As we already advanced in the introduction of this dissertation (see Chapter 1, Section 1.1), learning Dutch is not effortless for adult Spaniards, and when you ask them what the most difficult aspect of learning Dutch is, most of them will answer: “*la pronunciación*”, ‘the pronunciation’. We believe that the outcomes of this dissertation throw light on the specific pronunciation problems Spanish learners of Dutch have, as well as their sources. Such insights can help:

- 1) to propose pedagogical direction for phonological instruction in the Dutch L2 classroom;
- 2) to develop dedicated CAPT programs;
- 3) to create materials aimed at raising phonological awareness among Spanish learners.

Hopefully, valorization activities along these lines can maximize the societal impact of the findings in this dissertation, and can contribute to the personal and professional integration of my fellow Spanish-speaking peers in the Netherlands.

References

- Adank, P. (2003). *Vowel normalization: A perceptual-acoustic study of Dutch vowels*. Doctoral dissertation. Nijmegen, the Netherlands: Radboud University Nijmegen.
- Adank, P., Smits, R., & Van Hout, R. (2004a). A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America*, 116(5), 3099–3107.
- Adank, P., Van Hout, R., & Smits, R. (2004b). An acoustic description of the vowels of Northern and Southern Standard Dutch. *Journal of the Acoustical Society of America*, 116(3), 1729–1738.
- Akkermans, D., Harzing, A. W., & Van Witteloostuijn, A. (2010). Cultural accommodation and language priming: competitive versus cooperative behavior in a prisoner's dilemma game. *Management International Review*, 50(5), 559–583.
- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgements of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning*, 42(4), 529–555.
- Baker, W., & Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech*, 48(1), 1–27.
- Bassetti, B. (2006). Orthographic input and phonological representations in learners of Chinese as a foreign language. *Written Language and Literacy*, 9, 95–114.
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America*, 114(3), 1600–1610.
- Bent, T., Baesse-Berk, M., Borrie, S. A., & McKee, M. (2016). Individual differences in the perception of regional, nonnative, and disordered speech varieties. *Journal of the Acoustical Society of America*, 140(5), 3775–3786.
- Berck, P. (2017). *Memory-based text correction*. Doctoral dissertation. Nijmegen, the Netherlands: Radboud University Nijmegen.
- Best, C. T. (1995). A direct realist view of speech cross language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 171–206). Timonium, MD: York Press.
- Best, C. T., & Tyler, M. D. (2007). 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* (pp. 13–34). Amsterdam: John Benjamins.
- Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language*, 44(2), 235–249.
- Boersma, P., & Weenink, D. (2010). Praat: Doing phonetics by computer. Version 5.1.32. Retrieved from <http://www.fon.hum.uva.nl/praat/> (date last viewed 31/08/17).
- Bohn, O.-S. (1995). Cross language speech perception in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 279–304). Timonium, MD: York Press.
- Bongaerts, T. (2005). Introduction: Ultimate attainment and the critical period hypothesis for second language acquisition. *IRAL*, 43, 259–267.

- Bongaerts, T., Planken, B., & Schils, E. (1995). Can late learners attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition* (pp. 30–50). Clevedon, UK: Multilingual Matters.
- Booij, G. (1995). *The phonology of Dutch*. Oxford, United Kingdom: Clarendon Press.
- Botma, B., & Van Oostendorp, M. (2012). À propos of the Dutch vowel system 21 years on, 22 years on. In Botma, B., & Noske, R. (Eds.), *Phonological explorations: Empirical, theoretical and diachronic issues* (pp. 135–154). Berlin: Mouton de Gruyter.
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106, 707–729.
- Brennan, E., & Brennan, J. (1981). Accent scaling and language attitudes: Reactions to Mexican American English speech. *Language and Speech*, 24(3), 207–221.
- Brown, B., Giles, H., & Thakerar, J. (1985). Speaker evaluations as a function of speech rate, accent and context. *Language and Communication*, 5, 207–220.
- Burgos, P., Cucchiari, C., Van Hout, R., & Strik, H. (2013). Pronunciation errors by Spanish learners of Dutch: A data-driven study for ASR-based pronunciation training. *Proceedings of Interspeech 2013*, Lyon, France, 2385–2389.
- Burgos, P., Cucchiari, C., Van Hout, R., & Strik, H. (2014a). Phonology acquisition in Spanish learners of Dutch: Error patterns in pronunciation. *Language Sciences*, 41, 129–142.
- Burgos, P., Jani, M., Cucchiari, C., Van Hout, R., & Strik, H. (2014b). Dutch vowel production by Spanish learners: Duration and spectral features. *Proceedings of Interspeech 2014*, Singapore, 529–533.
- Burgos, P., Sanders, E., Cucchiari, C., Van Hout, R., & Strik, H. (2015). *Auris Populi*: Crowdsourced native transcriptions of Dutch vowels spoken by adult Spanish learners. *Proceedings of Interspeech 2015*, Dresden, Germany, 2819–2823.
- Burgos, P., Jani, M., Van Hout, R., Cucchiari, C., & Strik, H. Spanish-accented Dutch vowel productions: Duration and spectral features (submitted a).
- Burgos, P., Sanders, E., Van Hout, R., Cucchiari, C., & Strik, H. *Auris Populi*: Assessing the intelligibility of L2 speech (submitted b).
- Burgos, P., Van Hout, R., & Planken, B. Matching acoustical properties and native perceptual assessments of L2 speech (submitted c).
- Cargile, A. (2000). Evaluations of employability: Does accent always matter? *Journal of Employment Counseling*, 37, 165–177.
- Cargile, A., & Giles, H. (1998). Language attitude towards varieties of English: An American-Japanese context. *Journal of Applied Communication Research*, 26, 338–356.
- Carlisle, R. S. (1991). The influence of environment on vowel epenthesis in Spanish/English interphonology. *Applied Linguistics*, 12(1), 76–95.
- Carlson, H. K., & McHenry, M. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling*, 43, 70–83.
- Celce-Murcia, M., Brinton, D.M., & Goodwin, J. M. (1996). *Teaching pronunciation*. Cambridge: University Press.

- Centraal Bureau voor de Statistic*, Dutch National Institute of Statistics. Retrieved from <http://www.cbs.nl/en-gb/menu/themas/gezondheid-welzijn/publicaties/artikelen/archief/2008/2008-2367-wm.htm> (date last viewed 31/08/17).
- Cervera, T., Miralles, J. L., & González-Álvarez, J. (2001). Acoustical analysis of Spanish vowel produced by laryngectomized subjects. *Journal of Speech, Language and Hearing Research*, 44, 988–996.
- Chládková, K., Escudero, P., & Boersma, P. (2011). Context-specific acoustic differences between Peruvian and Iberian vowels. *Journal of the Acoustical Society of America*, 130(1), 416–428.
- Chládková, K., & Podlipský, V. J. (2011). Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners. *Journal of the Acoustical Society of America Express Letters*, 130(4), 186–192.
- Chládková, K., Podlipský, V. J., & Chionidou, A. (2017). Perceptual adaptation of vowels generalizes across the phonology and does not require local context. *Journal of Experimental Psychology*, 43(2), 414–427.
- Clarke, C. M., & Luce, P. A. (2005). Perceptual adaptation to speaker characteristics: VOT boundaries in stop voicing categorization. *Proceedings of the ISCA Workshop on Plasticity in Speech Perception (PSP 2005)*. London: University College London, 23–26.
- Clarke, C. M., & Garret, M. F. (2004). Rapid adaptation to foreign accented English. *Journal of the Acoustical Society of America*, 116(6), 3647–3658.
- Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston & M. Beckmann (Eds.), *Papers in Laboratory Phonology 1: Between the Grammar and the Physics of Speech* (pp. 283–333). Cambridge: Cambridge University Press.
- Cohen, A., Slis, I. H., & 't Hart, J. (1963). Perceptual tolerance of isolated Dutch vowels. *Phonetica*, 9, 65–78.
- Cohen, A., Slis, I. H., & 't Hart, J. (1967). On tolerance and intolerance in vowel perception. *Phonetica*, 16, 65–70.
- Colonia española de Eindhoven*, 'Spanish working community in Eindhoven'. Retrieved from <http://www.emigracioneindhoven.dse.nl/> (date last viewed 31/08/17).
- Council of Europe. Common European Framework of Reference for Languages (CEFR). Retrieved from http://www.coe.int/t/dg4/linguistic/cadre1_en.asp (date last viewed 31/08/17).
- Council of Europe. CEFR Self-Assessment Grid. Retrieved from <https://rm.coe.int/168045bb52> (date last viewed 31/08/17).
- Cucchiari, C. (1996). Assessing transcription agreement: Methodological aspects. *Clinical Linguistics and Phonetics*, 10(2), 131–155.
- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.
- De Angelis, G. (2007). *Third or Additional Language Acquisition*. Clevedon, United Kingdom: Multilingual Matters.
- Derwing, T. M. (2008). Curriculum issues in teaching pronunciation to second language learners. In J.G. Hansen Edwards, & M.L. Zampini (Eds.). *Phonology and second language acquisition* (pp. 347–369). Amsterdam: John Benjamins.

- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility and comprehensibility. Evidence from four L1s. *SSLA. Studies in Second Language Acquisition*, 20, 1–16.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching. A research based approach. *TESOL Quarterly*, 39(3), 379–397.
- Derwing, T. M., & Munro, M. J. (2015). A prospectus for pronunciation research in the 21st century. *Journal of Second Language Pronunciation*, 1(1), 11–42.
- Deprez-Sims, A.-S., Morris, S. B. (2010). Accents in the workplace: Their effects during a job interview. *International Journal of Psychology*, 45(6) 417–426.
- Ehsani, F., & Knodt, E. (1998). Speech Technology in computer-added language learning: Strengths and limitations of a new CALL paradigm. *Language Learning & Technology*, 2(1), 54–73.
- Erdener, D. V., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, 55, 191–228.
- Ernestus, M., Kouwenhoven, H., & Van Mulken, M. (2017). The direct and indirect effects of the phonotactic constraints in the listener's native language on the comprehension of reduced and unreduced word pronunciation variants in a foreign language. *Journal of Phonetics*, 62, 50–64.
- Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Doctoral dissertation. Utrecht, the Netherlands: Utrecht University.
- Escudero, P. (2006). The phonological and phonetic development of new vowel contrast in Spanish learners of English. In Baptista, B. O. & Watkins, M. A. (Eds.), *English with a Latin beat: Studies in Portuguese/Spanish-English interphonology. Studies in bilingualism* (pp. 149–161). Amsterdam: John Benjamins.
- Escudero, P. (2015). Orthography plays a limited role when learning the phonological forms of new words: The case of Spanish and English learners of novel Dutch words. *Applied Psycholinguistics*, 36, 7–22.
- Escudero, P., Benders, T., & Lipski, S. (2009). 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.
- Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, 26(4), 551–585.
- Escudero, P., Simon, E., & Mulak, K. (2014). Learning words in a new language: Orthography doesn't always help. *Bilingualism: Language and Cognition*, 17(2), 384–395.
- Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech*, 53(3), 343–365.
- Escudero, P., & Williams, D. (2011). Perceptual Assimilation of Dutch vowels by Peruvian Spanish listeners. *Journal of the Acoustical Society of America*, 129(1), 1–7.
- Escudero, P., & Williams, D. (2012). 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.

- Eskenazi, M. (1999). Using automatic speech processing for foreign language pronunciation tutoring: Some issues and a prototype. *Language Learning & Technology*, 2, 62–76.
- Eskenazi, M. (2009). An overview of spoken language technology for education. *Speech Communication*, 51, 832–844.
- Eskenazi, M., Levow, G. A., Meng, H., Parent, G., & Suendermann, D. (2013). *Crowdsourcing for Speech Processing: Applications to Data Collection, Transcription and Assessment*. Hoboken: Wiley Publishing.
- Fayer, J. M., & Krasinski, E. (1987). Native and nonnative judgments of intelligibility and irritation. *Language Learning*, 37, 313–326.
- Flege, J. E. (1991). Orthographic evidence for the perceptual identification of vowels in Spanish and English. *Quarterly Journal of Experimental Psychology*, 40, 701–31.
- Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 233–277). Timonium, MD: York Press.
- Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the Critical Period Hypothesis* (pp. 101–131). Hillsdale, NJ: Erlbaum.
- Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer and N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–355). Berlin: Mouton de Gruyter.
- Flege, J. E. (2005a). “Origins and development of the Speech Learning Model”, lecture presented at the 1st ASA Workshop on L2 Speech Learning, Simon Fraser University, Vancouver, BC, Canada, 14–15 April 2005. Retrieved from <http://www.jimflege.com/> (date last viewed 31/08/17).
- Flege, J. E. (2005b). “Evidence for plasticity in studies examining second language speech acquisition” (“Doom”(no plasticity) hypothesis), lecture presented at the ISCA Workshop on Plasticity in Speech Perception (PSP2005), London, UK, 15–17 June 2005. Retrieved from <http://www.jimflege.com/> (date last viewed 31/08/17).
- Flege, J. E. (2009). Give input a chance! In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 175–190). Bristol, UK: Multilingual Matters.
- Flege, J. E. (2012). “The role of input in second language (L2) speech learning”, lecture presented at the VIth International Conference on Native and Non-native Accents of English, Łódź, Poland, 6–8 December 2012. Retrieved from <http://www.jimflege.com/> (date last viewed 31/08/17).
- Flege, J., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Tsukada, K. (2006). Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics*, 33, 153–175.
- Flege, J. E., Bohn, O.-S. (1989). An instrumental study of vowel reduction and stress placement in Spanish accented English. *SSLA. Studies in Second Language Acquisition*, 11, 35–62.
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). The effect of experience on nonnative subjects’ production and perception of English vowels. *Journal of Phonetics*, 25, 437–470.

- Flege, J. E., & Eefting, W. (1987). The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics*, 15, 67–83.
- Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, 25, 169–186.
- Flege, J. E., Schirru, C., & MacKay, I. R. A. (2003). Interaction between the native and the second language phonetic subsystems. *Speech Communication*, 40, 467–491.
- Flege, J. E., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, 41, 78–104.
- Gass, S., & Varonis, E. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, 34, 65–89.
- Geva, E., & Wang, M. (2001). The development of basic reading skills in children: A cross language perspective. *Annual Review of Applied Linguistics*, 21, 182–204.
- Goeman, T. (2002). Perception of dialect distance: Standard and dialect in relation to new data on Dutch varieties. In D. Long, & D. R. Preston (Eds.), *Handbook of perceptual dialectology* (pp. 135–149). Amsterdam, the Netherlands: John Benjamins.
- Gottfried, T. (2008). Music and language learning. Effects of musical training on learning L2 speech contrasts. In O.-S. Bohn and M. Munro (Eds.), *Language experience in second language speech learning* (pp. 221–237). Amsterdam: John Benjamins.
- Goudbeek, M., Cutler, A., & Smits, R. (2008). Supervised and unsupervised learning of multidimensionally varying non-native speech categories. *Speech Communication*, 50, 109–125.
- Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals. *Phonetica*, 60, 98–128.
- Gussenhoven, C. (1999). Vowel duration, syllable quantity and stress in Dutch. Centre for Language Studies, Nijmegen, the Netherlands.
- Gussenhoven, C. & Jacobs, H. (2011). *Understanding phonology*. (3rd ed.). London: Hodder Education.
- Hammond, R. M. (2001). *The sounds of Spanish: Analysis and application*. Sommerville, MA: Cascadilla Press.
- Harzing, A.-W., Köster, K., & Magner, U. (2011). Babel in business: The language barrier and its solutions in the HQ-subsidiary relationship. *Journal of World Business*, 46, 279–287.
- Harzing, A.-W., & Pudelko, M. (2014). Hablas vielleicht un peu la mia language? A comprehensive overview of the role of language differences in headquarters–subsidiary communication. *The International Journal of Human Resource Management*, 25(5), 696–717.
- Hazan, V., Sennema, A., & Faulkner, A. (2002). Audiovisual perception in L2 learners. *Proceedings of the International Conference for Spoken Language Processing 2002*, Denver, USA, 1685–1688.
- Hualde, J. I. (2005). *The sounds of Spanish*. Cambridge, United Kingdom: Cambridge University Press.
- Instituto Nacional de Estadística*, Spanish National Institute of Statistics, Retrieved from <http://www.ine.es/jaxi/tabla.do> (date last viewed 05/05/13).

- Iverson, P., & Evans, B. G. (2007). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movements, and duration. *Journal of the Acoustical Society of America*, 122(5), 2842–2854.
- Jani, M., Cucchiarini, C., Van Hout, R., & Strik, H. (2015). Confusability in L2 vowels: Analyzing the role of different features. *Proceedings of Interspeech 2015*, Dresden, Germany, 1715–1719.
- Jones, R. H. (1997). Beyond ‘Listen and Repeat’: Pronunciation teaching materials and theories of Second Language Acquisition. *System*, 25, 103–112.
- Julkowska, I. A. & Cebrian, J. (2015). Effect of listener factors and stimulus properties on the intelligibility, comprehensibility and accentedness of L2 speech. *Journal of Second Language Pronunciation*, 1(2), 211–237.
- Kankaanranta, A., & Planken, B. (2010). BELF competence as business knowledge of internationally operating business professionals. *Journal of Business Communication*, 47(4), 380–407.
- Kendall, T., & Thomas, E. R. (2010). Vowels: Vowel Manipulation, Normalization, and Plotting in R. R package, version 1.1. Retrieved from <http://ncslaap.lib.ncsu.edu/tools/norm/> (date last viewed 05/05/13).
- Kennedy, S. & Trofimovich, P. (2008) Intelligibility, comprehensibility, and accentedness: The role of listener experience and semantic context. *The Canadian Modern Language Review*, 64(3), 459–489.
- Kingston, P. (1996). Bridging the language gap through international networking. In M. Berger (Ed.), *Cross-cultural team building: Guidelines for effective communication and negotiation* (pp. 58–70). London: McGraw-Hill.
- Koet, A. (2007). *Polder English in Dutch ears. Empirical studies on the evaluation of the pronunciation of English as a foreign language*. Doctoral Dissertation. Amsterdam, the Netherlands: University of Amsterdam.
- Kolb, A., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning & Education*, 4(2), 193–212.
- König, E., & Van der Auwera, J. (1994). *The Germanic languages*. London: Routledge.
- Kouwenhoven, H. (2016). *Situational variation in non-native communication: Studies into register variation, discourse management and pronunciation in Spanish English*. Doctoral dissertation. Nijmegen, the Netherlands: Radboud University Nijmegen.
- Kunath, S. A., & Weinberger, S. H. (2010). The wisdom of the crowd's ear: speech accent rating and annotation with Amazon Mechanical Turk. *Proceedings of CSLDAMT '10, Association for Computational Linguistics*, 168–170.
- Lapesa, R. (1981). *Historia de la lengua española*. (11th ed.). Madrid: Gredos.
- Lev-Ari, S., & Keysar B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46, 1093–1096.
- Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy and fluency in second language acquisition. *Applied Linguistics*, 30, 579–589.
- Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.

- Lippi-Green, R. (1997). *English with an accent: Language, ideology, and discrimination in the United States*. New York, NY: Routledge.
- Lobanov, B. M. (1971). Classification of Russian vowels spoken by different speakers. *Journal of the Acoustical Society of America*, 49, 606–608.
- Long, M. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12, 251–285.
- Lybeck, K. (2002). Cultural identification and second language pronunciation of Americans in Norway. *Modern Language Journal*, 86, 174–191.
- Magen, H. S. (1998). The perception of foreign-accented speech. *Journal of Phonetics*, 26, 381–400.
- Mai, R., & Hoffmann, S. (2014). Accents in Business Communication: An integrative model and propositions for future research. *Journal of Consumer Psychology*, 24(1), 137–158.
- Major, R. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum.
- Masgoret, A., & Gardner, R. (2003). Attitudes, motivation and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, 53, 123–163.
- Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism: Language and Cognition*, 13, 279–297.
- McAllister, R., Flege, J.E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30, 229–258.
- Morrison, G. S. (2006). *L1 & L2 production and perception of English and Spanish vowels: A statistical modelling approach*. Doctoral dissertation. Alberta, Canada: University of Alberta.
- Morrison, G. S. (2008). L1-Spanish speakers' acquisition of the English /i/-/ɪ/ contrast: duration-based perception is not the initial developmental stage. *Language and Speech*, 51, 285–315.
- Morrison, G. S. (2009). L1-Spanish speakers' acquisition of the English /i/-/ɪ/ contrast. II. Perception of vowel inherent spectral change. *Language and Speech*, 52, 437–462.
- Moyer, A. (1999). Ultimate attainment in L2 phonology. The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, 21, 81–108.
- Moyer, A. (2004). *Age, accent and experience in second language acquisition. An integrated approach to critical period inquiry*. Clevedon, UK: Multilingual Matters.
- Moyer, A. (2007). Do language attitudes determine accent? A study of bilinguals in the US. *Journal of Multilingual and Multicultural Development*, 28, 1–17.
- Moyer, A. (2013). *Foreign accent. The phenomenon of non-native speech*. Cambridge, UK: Cambridge University Press.
- Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, 36(1), 39–66.

- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97.
- Munro, M. J., & Derwing, T. M. (2015). A prospectus for pronunciation research in the 21st century: A point of view. *Journal of Second Language Pronunciation*, 1(1), 11–42.
- Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *SSLA. Studies in Second Language Acquisition*, 28, 111–131.
- Navarro Tomás, T. (2004). *Manual de pronunciación española*. (28th ed.). Madrid, Spain: Editorial CSIC-CSIC Press.
- Neri, A., Cucchiari, C., & Strik, H. (2006). Selecting segmental errors in L2 Dutch for optimal pronunciation training. *IRAL - International Review of Applied Linguistics in Language Teaching*, 44, 357–404.
- Neri, A., Cucchiari, C., & Strik, H. (2008). The effectiveness of computer-based corrective feedback for improving segmental quality in L2-Dutch. *ReCALL*, 20(2), 225–243.
- Neri, A., Cucchiari, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in Computer Assisted Pronunciation Training. *Computer Assisted Language Learning*, 15(5), 441–467.
- Nespor, M., Shukla, M., & Mehler, J. (2011). Stress-timed vs. syllable-timed languages. In M. van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology* (pp. 1147–1159). Malden: Wiley Blackwell.
- Nooteboom, S. G., & Doodeman, G. J. N. (1980). Production and perception of vowel length in spoken sentences. *Journal of the Acoustical Society of America*, 67, 276–287.
- Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, 25, 26–45.
- Oostdijk, H. J. (2002). The design of the Spoken Dutch Corpus. In P. Peters, P. Collins, & A. Smith (Eds.), *New frontiers of corpus research* (pp. 105–112). Amsterdam, the Netherlands: Rodopi.
- Ortega-Llebaria, M., Faulkner, A., & Hazan, V. (2001). Auditory-visual L2 speech perception: Effects of visual cues and acoustic-phonetic context for Spanish learners of English. *Proceedings of the Auditory-Visual Speech Processing International Conference 2001*, Aalborg, Denmark, 149–154.
- Parent G., & Eskenazi, M. (2011). Speaking to the Crowd: Looking at past achievements in using crowdsourcing for speech and predicting future challenges. *Proceedings of Interspeech 2011*, Florence, Italy, 3037–3040.
- Parker, S. G. (2002). *Quantifying the sonority hierarchy*. Doctoral dissertation. Amherst, MA, United States: University of Massachusetts Amherst.
- Parker, S. G. (2012). *The sonority controversy*. Berlin: De Gruyter Mouton, 2012
- Parnell, M., & Amerman, J. (1978). Maturation influences on the perception of coarticulatory effects. *Journal of Speech and Hearing Research*, 21, 682–701.
- Pavlenko, A., & Lantolf, J. (2000). Second language learning as participation and the (re)construction of selves. In J. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 155–178). Oxford, UK: Oxford University Press.

- Pennington, M.C. (1999). Computer-aided pronunciation pedagogy: Promise, limitations, directions. *Computer Assisted Language Learning*, 12, 427–440.
- Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29, 191–215.
- Quilis, A., & Fernández, J. A. (1985). *Curso de fonética y fonología española*. Madrid, Spain: Consejo Superior de Investigaciones Científicas.
- Rochet, B. L. (1995). Cross-language studies of speech perception and production in second-language learning. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research* (pp. 379–410). Timonium, MD: York Press.
- Rogers, C. L., & Dalby, J. M. (1996). Prediction of foreign-accented speech intelligibility from segmental contrast measures. *Journal of the Acoustical Society of America*, 100(4), 2725–2726.
- Rosenfelder, I., Fruehwald, J., Evanini, K., Seyfarth, S., Gorman, K., Prichard, H., & Yuan, J. (2014). FAVE 1.1.3. ZENODO. Retrieved from <http://fave.ling.upenn.edu/> (date last viewed 31/08/17).
- Sanders, E., Burgos, P., Cucchiarini, C., & Van Hout, R. (2016). *Palabras*: Crowdsourcing transcriptions of L2 speech. *Proceedings of the International Conference on Language Resources and Evaluation (LREC) 2016*, Portorož, Slovenia, 3186–3191.
- Schepens, J. J. (2015). *Bridging linguistic gaps: The effects of linguistic distance on the adult learnability of Dutch as an additional language*. Doctoral dissertation. Nijmegen, the Netherlands: Radboud University Nijmegen.
- Schepens, J. J., Van der Slik, F., & Van Hout, R. (2016). L1 and L2 distance effects in learning L3 Dutch. *Language Learning*, 66(1), 224–256.
- Scovel, T. (1988). *A Time to Speak. A Psycholinguistic Inquiry into the Critical Period for Human Speech*. New York: Newbury House.
- Silveira, R. (2007). Investigating the role of orthography in the acquisition of L2 pronunciation: A case study. *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*, 459–472.
- Sociaal en Cultureel Planbureau*, Dutch Social and Cultural Planning Office. Retrieved from <https://www.scp.nl/Zoekresultaten?searchbase=0&searchrange=10&searchpage=1&freetext=Summary+New+Spanish+migrants+in+the+Netherlands&submit=Zoeken> (date last viewed 31/08/17).
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39, 456–466.
- Strik, H., Truong, K., De Wet, F., & Cucchiarini, C. (2009). Comparing different approaches for automatic pronunciation error detection. *Speech Communication*, 51(10), 845–852.
- Suzuki, R., & Shimodaira, H. (2006). Pvcust: An R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*, 22(12), 1540–1542.
- Tenzer, H., & Pudelko, M. (2017). The influence of language differences on power dynamics in multinational teams. *Journal of World Business*, 52, 45–61.
- Tokuhamma-Espinosa, T. (2003). The relationship between musical ability and foreign languages. In T. Tokuhamma-Espinosa (Ed.), *The multilingual mind* (pp. 65–80). Westport, CT: Praeger.

- Trofimovich, P., Lightbown, P. M., Halter, R., & Song, H. (2009). Comprehension-based practice: The development of L2 pronunciation in a listening and reading program. *Studies in Second Language Acquisition*, 31, 609–639.
- Tropf, H. S. (1987). Sonority as a variability factor in second language phonology. In A. James & Leather, J. (Eds.), *Sound patterns in second language acquisition* (pp. 173–191). Providence, RI: Foris Publications.
- Tropf, H. S. (1983). *Variation in der Phonologie des ungesteuerten Zweitspracherwerbs*. Doctoral dissertation. Heidelberg, Germany: University of Heidelberg.
- Ulbrich, C. Phonetic drift: Effects of context and language experience in the acquisition of German vowel contrasts (in preparation).
- University of Cambridge Local Examinations Syndicate (UCLES) (2015). *Cambridge English. The Cambridge English Scale Explained*. Cambridge English Language Assessment. Retrieved from <http://www.cambridgeenglish.org/images/177867-the-methodology-behind-the-cambridge-english-scale.pdf> (date last viewed 31/08/17).
- Van den Doel, W. Z. (2006). *How friendly are the natives? An evaluation of native-speaker judgements of foreign-accented British and American English*. Doctoral Dissertation. Utrecht, the Netherlands: Utrecht University.
- Van der Harst, S. (2011). *The vowel space paradox: A sociophonetic study on Dutch*. Doctoral dissertation. Nijmegen, the Netherlands: Radboud University Nijmegen.
- Van der Harst, S., Van de Velde, H., & Van Hout, R. (2014). Variation in Standard Dutch vowels: The impact of formant measurement methods on identifying the speaker's regional origin. *Language, Variation and Change*, 26(2), 247–272.
- Van Doremalen, J., Cucchiari, C., & Strik, H. (2013). Automatic pronunciation error detection in non-native speech: The case of vowel errors in Dutch. *Journal of the Acoustical Society of America*, 134(2), 1336–1347.
- Van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, 6, 1–12.
- Van Maastricht, L., Krahmer, E., & Swerts, M. (2016a). Prominence patterns in a second language: Intonational transfer from Dutch to Spanish and vice versa. *Language Learning*, 66(1), 124–158.
- Van Maastricht, L., Krahmer, E., & Swerts, M. (2016b). Native speaker perceptions of (non-)native prominence patterns: Effects of deviance in pitch accent distributions on accentedness, comprehensibility, intelligibility, and nativeness. *Speech Communication*, 83, 21–33.
- Van Son, R. J. J. H., Binnenpoorte, D., Van den Heuvel, H., & Pols, L. C. W. (2001). The IFA corpus: a phonemically segmented Dutch open source speech database. *Proceedings of Eurospeech 2001*, Aalborg, Denmark, 2051–2054.
- Van Wijngaarden, S. J. (2001). Intelligibility of native and non-native Dutch speech. *Speech Communication*, 35, 103–113.
- Vijf eeuwen migratie*, 'five centuries of migration'. Retrieved from <http://www.vijfeeuwenmigratie.nl/term/Gastarbeiders> (date last viewed 31/08/17).
- Wade, T., Jongman, A., & Sereno, J. (2007). Effects of acoustic variability in the perceptual learning of non-native-accented speech sounds. *Phonetica*, 64, 122–144.

- Walley, A. C., & Flege, J. E. (2000). Effects of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics*, 27, 307–332.
- Wanrooij, K. (2015). Doctoral dissertation. Amsterdam, the Netherlands: University of Amsterdam.
- Wanrooij, K., Escudero, P., & Raijmakers, E. J. (2013). What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning. *Journal of Phonetics*, 41, 307–319.
- Wanrooij, K., & Boersma, P. (2013). Distributional training of speech sounds can be done with continuous distributions. *Journal of the Acoustical Society of America Express Letters*, 133(5), 398–404.
- Wells, J. (2004). SAMPA. Computer readable phonetic alphabet. Retrieved from <http://www.phon.ucl.ac.uk/home/sampa> (date last viewed 31/08/17).
- Wester, M., Kessens, J. M., Cucchiari, C., & Strik, H. (2001). Obtaining phonetic transcriptions: A comparison between expert listeners and a continuous speech recognizer. *Language and Speech*, 44(3), 377–403.
- Wik, P., & Escribano, D. L. (2009). Say 'Aaaaa' Interactive Vowel Practice for Second Language Learning. *Proceedings of SLATE* (pp. CD-rom). Warwickshire, U.K. Retrieved from <http://www.eee.bham.ac.uk/SLaTE2009/papers/SLaTE2009-4.pdf> (date last viewed 31/08/17).
- Wik, P., Hincks, R., & Hirschberg, J. (2009). Responses to Ville: A virtual language teacher for Swedish. *Proceedings of SLATE* (pp. CD-rom). Warwickshire, U.K. Retrieved from <http://www.eee.bham.ac.uk/SLaTE2009/papers/SLaTE2009-24.pdf> (date last viewed 31/08/17).
- Witt, S. M. (1999). *Use of speech recognition in computer-assisted language learning*. Doctoral dissertation. Cambridge, United Kingdom: University of Cambridge.
- Witteman, M. J., Weber, A., & McQueen, J. M. (2014). Tolerance for inconsistency in foreign-accented speech. *Psychonomic Bulletin & Review*, 21, 512–519.
- Yavaş, M., & Someillan, M. (2005). Patterns of acquisition of /s/-clusters in Spanish-English bilinguals. *Journal of Multilingual Communication Disorders*, 3(1), 50–55.
- Yeni-Komshian, G., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition*, 3(2), 131–149.
- Young-Scholten, M. (2002). Orthographic input in L2 phonological development. In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated view of language development - Papers in honour of Henning Wode* (pp. 263–279). Trier, Germany: Wissenschaftlicher Verlag Trier.
- Zec, D. (1995). Sonority constraints on syllable structure. *Phonology*, 12, 85–129.

Samenvatting (Dutch summary)

Globalisering en de toenemende mobiliteit hebben een enorm effect gehad op het onderwijs en de manier waarop mensen talen leren. De noodzaak om een buitenlandse of een tweede taal (L2) te leren is essentieel voor mensen die in een internationale context werken of in een L2 omgeving leven. Volwassen leerders worden geconfronteerd met een reeks van uitdagingen bij het verwerven van een L2. Het gaat om nieuwe morfologische paradigma's, syntactische structuren, lexicale elementen en fonologische kenmerken. Hoewel L2 sprekers vloeiend kunnen zijn, is het onwaarschijnlijk dat hun spraak met een accent – vaak door interferentie van hun moedertaal (L1) – zo verstaanbaar kan zijn als de spraak van native sprekers en dat kan van invloed zijn op de effectiviteit van de communicatie. Het is algemeen bekend dat volwassen L2 leerders veel moeite hebben om L2 spraakklanken tot in detail te verwerven en veel van hen behouden een buitenlands accent, zelfs nadat ze meerdere jaren in het gastland hebben doorgebracht. Het hebben van een buitenlands accent kan sociale gevolgen hebben. De competenties van volwassenen L2 leerders worden vaak beoordeeld op basis van hun buitenlandse accent, wat nadelig kan zijn voor carrièremogelijkheden, succesvolle interactie en sociale acceptatie.

Tien jaar geleden, na de economische crisis in 2008 en de sterke stijging van de Spaanse werkloosheid, kwamen vele nieuwe Spaanse migranten naar Nederland. Deze Spaanse migranten zijn mobiel, hoog opgeleid en spreken goed Engels. De meeste werken in de hightech- en zorgsector of zijn ingeschreven als studenten in het hoger onderwijs in Nederland. Meer specifiek, 65% van de Spaanse migranten die in Nederland werken, werken in vakspecialistische of technische banen, bijvoorbeeld als onderzoekers, docenten, ingenieurs, verpleegkundigen of IT-specialisten. De meeste hebben echter geen vaardigheid in het Nederlands wanneer ze in Nederland

aankomen. Terwijl ze aanvankelijk in het Engels uit de voeten kunnen, worden ze zich snel bewust van het belang van het spreken van het Nederlands, omdat het op het werk nodig is of omdat ze hun sociale interactie willen verbeteren.

Nederlands leren is moeilijk voor volwassen Spanjaarden en wanneer hun wordt gevraagd wat het moeilijkste aspect van het leren van het Nederlands is, zullen de meeste van hen waarschijnlijk antwoorden: “*la pronunciación*”, de uitspraak. Het hoofddoel van dit onderzoek is de analyses van de uitspraakproblemen van volwassen Spaanse leerders van het Nederlands en de mogelijke oorzaken ervan, alsook om erachter te komen hoe native Nederlandse luisteraars de Nederlandse uitspraak met een Spaans accent percipiëren wat betreft verstaanbaarheid. Het onderzoek richt zich op vier onderzoeksvragen:

OV₁: Wat zijn de meeste frequente uitspraakproblemen van volwassen Spaanse leerders van het Nederlands en wat zijn de mogelijke oorzaken van deze uitspraakproblemen?

OV₂: Komen de Nederlandse klinkers uitgesproken door volwassen Spaanse leerders akoestisch overeen met de klinkers die door native Nederlandse sprekers zijn uitgesproken?

OV₃: Zijn de Nederlandse klinkers door Spaanse leerders van het Nederlands uitgesproken verstaanbaar voor niet-deskundige native Nederlandse luisteraars?

OV₄: *Komen de akoestische kenmerken van de Nederlandse klinkers gesproken door volwassen Spaanse leerders van het Nederlands overeen met perceptuele beoordelingen van deze klinkers?*

Hoofdstuk 2 richt zich op het beantwoorden van OV₁, namelijk *Wat zijn de meeste frequente uitspraakproblemen van volwassen Spaanse leerders van het Nederlands, en wat zijn de mogelijke oorzaken van deze uitspraakproblemen?* Uit onze bevindingen blijkt dat onder de Spaanse leerders van het Nederlands klinkerfouten frequenter, persistenter en variabeler zijn dan fouten bij consonanten. Spaanse leerders lijken problemen te hebben met contrasten in klinkerlengte en klinkerhoogte en met geronde voorklinkers. Wat Nederlandse consonanten betreft treden er problemen in enkelvoudige fonemen op aan woordbegin en wordeinde, maar vooral in clusters, die verantwoordelijk zijn voor een groot aantal invoegingen, substituties en deleties. Deze uitspraakfouten zijn het gevolg van interferentie van de Spaanse fonologie en soms ook het Engels. Er zijn fonotactische beperkingen die van toepassing zijn in de moedertaal (d.w.z. open syllabestructuur (CV) in het Spaans tegen gesloten structuur (CVC) in het Nederlands; er zijn verschillen in de sonoriteitshierarchie bij de bepaling van toelaatbare consonantclusters). Ook zijn er uitspraakfouten als gevolg van orthografische interferentie voor zowel klinkers als consonanten.

Hoofdstuk 3 behandelt de productieve nauwkeurigheid van Nederlandse vocalen uitgesproken door Spaanse leerders. Hoofdstuk 3 heeft tot doel om OV₂ te beantwoorden, d.w.z. *Komen de Nederlandse klinkers uitgesproken door volwassen Spaanse leerders akoestisch overeen met klinkers die door native Nederlandse sprekers zijn uitgesproken?* De Nederlandse klinkers van volwassen Spaanse leerders zijn akoestisch geanalyseerd en de realisaties van de leerders zijn vergeleken met de realisaties van Nederlandse sprekers. Het

spraakmateriaal bestond per Spaanse spreker uit een set van 29 monosyllabische Nederlandse woorden die alle Nederlandse klinkers in beklemtoonde positie bevatten. Uit onze bevindingen blijkt dat de realisaties van de leerders niet overeenkomen met die van native speakers voor duur en spectrale waarden. Dit lijkt wederom toe te schrijven aan L1-verankerings-effecten. Het produceren van Nederlandse monoftongen is moeilijk voor Spaanse leerders, vooral wanneer klinkercontrasten subtiele spectrale verschillen weerspiegelen. Daarom gebruiken ze duur vaak foutief om dergelijke contrasten alsnog te realiseren. Daarentegen bleken de Spaanse leerders succesvol te zijn in het maken van het kort/lang onderscheid en in het produceren van de lange midden-vocalen en de diftongen van het Nederlands. Opmerkelijk genoeg waren meerdere leerders ook in staat om een nieuwe klinkercategorie te creëren, nl. een geronde voorklinker.

Hoofdstukken 4 en 5 onderzoeken de perceptie van het Nederlands gesproken met een Spaans accent door native Nederlandse luisteraars wat verstaanbaarheid betreft. Doel is de beantwoording van OV₃: *Zijn de Nederlandse klinkers door Spaanse leerders van het Nederlands uitgesproken verstaanbaar voor niet-deskundige native Nederlandse luisteraars?* Hoofdstuk 4 beschrijft een crowdsourcing studie waarin dezelfde Nederlandse monosyllabische woorden die door Spaanse leerders zijn uitgesproken en akoestisch zijn geanalyseerd in hoofdstuk 3, zijn gebruikt als spraakstimuli. Native Nederlandse luisteraars hebben deze spraakstimuli orthografisch getranscribeerd. Een overeenkomst tussen de klinker die door een native luisteraar is getranscribeerd en de canonieke (doel) vorm van dezelfde klinker laat zien dat de uitspraak van de Spaanse leerder voor native Nederlandse luisteraars verstaanbaar is. Het doel van de crowdsourcing studie was om te onderzoeken hoe de *auris populi*, het oor van het volk, omgaat met afwijkende realisaties van L2 klinkers. Uit de resultaten blijkt dat Nederlandse klinkers

die door Spaanse leerders zijn uitgesproken, op uiteenlopende manieren zijn getranscribeerd door native luisteraars. De transcripties van deze luisteraars bevestigen bevindingen van eerdere onderzoeken op basis van annotaties door deskundigen van de Spaanse leerders, namelijk dat de vijf Spaanse klinkers fungeren als 'attractors' voor Nederlandse klinkers. In het algemeen stemmen de resultaten ook overeen met de uitkomsten van de akoestische metingen. Een bonus van ons onderzoek was dat we konden laten zien dat de *auris populi*-methodologie een waardevol instrument is om L2-spraaktranscripties te verzamelen.

Hoofdstuk 5 onderzoekt eveneens hoe Nederlandse klinkers die door Spaanse leerders zijn uitgesproken door een diverse en grote groep van niet-deskundige native Nederlandse luisteraars worden waargenomen, maar met een andere, meer gecontroleerde steekproefmethode. Er is een sneeuwbalsteekproef gebruikt, die bestaat uit het werven van een groot aantal personen uit de sociale netwerken van een kleine startgroep van individuen. De resultaten laten opnieuw zien dat Nederlandse klinkers die door Spaanse leerders zijn uitgesproken vaak verschillend van hun canonieke vormen werden getranscribeerd. De resultaten consolideren eerdere bevindingen over de verstaanbaarheid van Nederlands met een Spaans accent. De gevonden klinkerverwarringspatronen stemmen overeen met de crowdsourced steekproef wat wederom het nut van “het oor van het volk” voor toekomstig L2-spraakonderzoek ondersteunt. Daarnaast werden er aanwijzingen gevonden voor perceptuele aanpassing, waardoor native luisteraars hun foneemgrenzen aanpassen aan de mix van niet-native en native spraak waaraan ze worden blootgesteld.

Hoofdstuk 6 richt zich op OV₄: *Komen de akoestische kenmerken van de Nederlandse klinkers gesproken door volwassen Spaanse leerders van het Nederlands overeen met perceptuele beoordelingen van deze klinkers?*

Statistische klinkerclassificaties op basis van akoestische kenmerken (akoestische data uit hoofdstuk 3) werden vergeleken met classificaties verkregen op basis van de oordelen van native Nederlandse luisteraars (luisteraar data uit hoofdstuk 5). Beide soorten classificaties bleken te worden beïnvloed door de specifieke set klinkers die als stimuli werden opgenomen, een effect dat herleid kon worden tot de grote variabiliteit in de klinkerrealisaties van de Spaanse leerders. Hoewel er sprake was van uitgesproken overeenkomsten tussen de twee soorten classificaties, werden ook verschuivingen gevonden binnen en tussen productie en perceptie, afhankelijk van de klinkerstimuli en de klinkerkenmerken. We onderzochten de variabiliteit tussen Spaanse leerders verder door individuele patronen in de productie en perceptie data te onderzoeken en deze te koppelen aan het taalvaardigheidsniveau en de meertalige achtergrond van de leerders. Onze resultaten voor individuele verschillen tussen Spaanse leerders, zowel voor de akoestische als de luisteraar data, bevestigen eerdere bevindingen (zie Burgos et al., 2014a; Flege, Frieda, & Nozawa, 1997; Munro, 1993; Yeni-Komshian, Flege, & Liu, 2000) door te laten zien dat het verwerven van de L2-fonologie niet altijd gelijk oploopt met de algemene taalvaardigheid in een tweede taal. We trekken de conclusie dat integratie van productie- en perceptiedata waardevolle inzichten oplevert in de rol van verschillende kenmerken in het leren van een tweede taal bij volwassen leerders en in hoe de kenmerken van klinkerrealisaties inwerken op de manier waarop L2-spraak wordt waargenomen. Een tweede conclusie is dat adaptieve mechanismen die nuttig zijn om te kunnen omgaan met variabiliteit van niet-native klinkerstimuli, een rol spelen in zowel statistische klinkerclassificaties (productie) als menselijke klankherkenning (perceptie). Tenslotte wijzen onze bevindingen erop dat variatie in het verwerven van de L2-fonologie extreem complexe vormen kan aannemen. Variabiliteit komt op verschillende niveaus voor, binnen en tussen

leerders met betrekking tot segmentale deviaties per target klinker, en binnen en tussen leerders met betrekking tot de onderliggende kenmerken (klinkerhoogte, lengte, ronding en diftongering) zoals die worden toegepast door leerders om Nederlandse klinkers te realiseren.

Dit onderzoek heeft antwoorden gegeven op de vier onderzoeksvragen. De drie meest relevante problemen die voortvloeien uit de onderzoeksresultaten, namelijk variabiliteit in individuele leerpaden, het optreden van adaptieve mechanismen (zowel in statistische classificatie als menselijke spraakherkenning) en de relatie tussen taalvaardigheid en uitspraakprestatie in de L2, worden hieronder besproken.

Individuele variabiliteit

De 28 volwassen Spaanse leerders van het Nederlands laten grote variatie zien in de manier waarop zij hun Nederlandse klinkers realiseren. Leerders maken niet allemaal hetzelfde onderscheid en niet alle onderscheidingen worden in dezelfde mate gemaakt. Onze resultaten geven aan dat de variabiliteit bij het verwerven van L2-klanken extreem ingewikkeld is en dat het op verschillende niveaus optreedt. Meer specifiek is er sprake van grote variatie binnen leerders, zowel in hun segmentale deviaties (zie Wade, Jongman, & Sereno, 2007) als in de manier waarop ze verschillende kenmerken gebruiken (klinkerhoogte, klinklengte, ronding en diftongering). Dit geldt voor alle Nederlandse doelklinkers. Ook is er een breed scala aan variabiliteit tussen leerders in hun vermogen om de Nederlandse target klinkers nauwkeurig te realiseren. De variabiliteit binnen en tussen leerders leidt tot duidelijke patronen van klinkerverwarringen per doelklinker (zie Bent et al., 2016; Mayr & Escudero, 2010). Leerders maken op uiteenlopende en ook afwijkende wijze gebruik van onderliggende kenmerken, om Nederlandse klinkers te realiseren. Variabiliteit wordt ook waargenomen in welke klinkercontrasten

het best worden geleerd. Met andere woorden, er is grote variatie in de paden die de leerders volgen in hun inspanningen om een nauwkeurige uitspraak van Nederlandse doelklinkers te bereiken.

Adaptieve mechanismen (statistische classificatie en menselijke klankherkenning)

De variabiliteit inherent aan de productie van L2-leerders impliceert dat statistische classificatie en menselijke klankherkenning zich moeten aanpassen aan verschillende uitspraakvarianten zowel binnen als tussen leerders om de beoogde betekenis te kunnen waarnemen. Onze uitkomsten leverden aanwijzingen voor het optreden van adaptieve mechanismen, zowel in de statistische klinkerclassificaties van de akoestische data als in de menselijke klinkerkenning, afhankelijk van de betrokken klinkersets (zie hoofdstuk 6). Patronen van klinkerconfusies in de statistische klinkerclassificaties van de native data gaven aan dat de statistische classifier gevoelig is en categoriegrenzen aanpast op grond van de aanwezige niet-native spraaksteekproeven. De classificatie bleek betere uitkomsten op te leveren wanneer slechts telkens één leerder aan het native spraakmateriaal werd toegevoegd in plaats van de hele verzameling van leerders (zie hoofdstuk 6). De input van grotere hoeveelheden niet-native data met een grote variatie in klinkerfouten lijkt te leiden tot grensverschuivingen omdat de statistische classifier moet zien om te gaan met “foutieve” data die aanzienlijk verschillen van de "schone" native data (zie Berck (2017) die laat zien hoe algoritmen voor machine learning worden beïnvloed door het invoeren van fouten in taaldata).

Soortgelijke adaptieve mechanismen werden waargenomen in de menselijke klankherkenning (zie hoofdstukken 5 en 6). De variabiliteit inherent aan de productie van L2-leerders impliceert dat native luisteraars zich

moeten aanpassen aan verschillende uitspraken van leerders. Om woorden met klankafwijkingen te herkennen moeten luisteraars deze klanken toewijzen aan foneemcategorieën (zie hoofdstukken 5 en 6). De native luisteraars moeten foneemgrenzen aanpassen om niet-native realisaties te classificeren op grond van hun native foneemcategorieën (zie Bent et al., 2016; Cutler, 2012; Bradlow & Bent, 2008). Als gevolg hiervan kunnen native luisteraars hun categoriegrenzen tijdelijk aanpassen of verplaatsen ten gunste van de dubbelzinnige fonemen in de niet-native steekproeven. Onze resultaten leveren bewijs voor wat Cutler (2012: 375) beschrijft als "de plasticiteit in de perceptie van volwassen native luisteraars". Perceptuele aanpassing van de native luisteraars kan een grensverschuiving tussen categorieën veroorzaken wanneer zij blootgesteld worden aan een mengsel van niet-native (cf. Clarke & Luce, 2005) en native spraak. Onze uitkomsten geven aan dat het native menselijke oor in staat is om een groot aantal variaties in het spraaksignaal van niet-native spraak te verwerken.

L2 taalvaardigheid en uitspraak

Onze bevindingen inzake individuele verschillen tussen de 28 volwassen Spaanse leerders van het Nederlands, zowel voor de akoestische als de luisteraar data, wijzen erop dat de uitspraakvaardigheid niet altijd overeenkomt met de algemene taalvaardigheid in een tweede taal (zie hoofdstuk 6, zie ook hoofdstuk 2). Een hoog algemeen vaardigheidsniveau in het Nederlands garandeert geen succes bij het verwerven van een native uitspraak in het Nederlands (zie hoofdstuk 6). Factoren die verband houden met algemene taalvaardigheid in een tweede taal zijn verblijfsduur en een intensief gebruik van de L2. Ook eerdere studies hebben evenwel laten zien dat deze factoren geen sterk effect hebben op de uitspraaknauwkeurigheid van

de L2 (zie Flege, Frieda, & Nozawa, 1997; Munro, 1993; Yeni-Komshian, Flege, & Liu, 2000).

Opmerkelijk genoeg worden de fonologische criteria waaraan leerders moeten voldoen om een bepaald taalvaardigheidsniveau in de L2 te bereiken niet specifiek beschreven in de huidige standaardmodellen van L2-leren (zie ook Moyer (2013) voor een discussie over accent binnen de huidige standaardmodellen). Dit gebrek aan specificiteit lijkt erop te wijzen dat fonologie niet als een essentieel vaardigheidsdomein wordt beschouwd in de huidige standaardmodellen van taalvaardigheid, zoals de CEFR. De bevindingen van dit onderzoek laten zien dat L2-fonologieverwerving extreem essentieel is in de verstaanbaarheid van L2-spraak en dus ook wezenlijk voor succesvolle communicatie in een L2-omgeving. De bevindingen van dit onderzoek illustreren eveneens de complexiteit van individuele variabiliteit. Het is verleidelijk om te concluderen dat de huidige standaardmodellen geen specifiekere beschrijving van fonologische controle per taalvaardigheidsniveau presenteren omdat uitspraakvaardigheid moeilijk is te definiëren gegeven de grote variabiliteit tussen L2-leerders.

De bevindingen uit dit onderzoek zullen hopelijk bijdragen aan de ontwikkeling van specifieke leermiddelen voor moedertaalsprekers van het Spaans die hun uitspraak van het Nederlands willen verbeteren. Zoals we hebben vastgesteld, verloopt het leren van het Nederlands niet moeiteloos voor volwassen Spanjaarden. De uitkomsten van deze proefschrift werpen licht op de specifieke uitspraakproblemen die Spaanse leerders van het Nederlands hebben evenals op de onderliggende oorzaken. Deze inzichten kunnen helpen:

- 1) om concrete pedagogische richtlijnen in fonologische instructies voor Nederlands als tweede taal voor te stellen;

- 2) om toegewijde CAPT-programma's te ontwikkelen;
- 3) om onderwijsmateriaal te ontwikkelen dat gericht is op het versterken van het fonologisch bewustzijn van Spaanse leerders.

Hopelijk kunnen valorisatieactiviteiten langs deze lijnen de maatschappelijke impact van de bevindingen in dit proefschrift versterken en bijdragen aan de persoonlijke en professionele integratie van mijn Spaanse collega's in Nederland.

Resumen (Spanish summary)

La globalización y la creciente movilidad ha tenido un enorme impacto en la educación y en la manera en la que se aprenden idiomas extranjeros. La necesidad de aprender un idioma extranjero o segunda lengua (L2) de una manera efectiva es esencial para poder operar en un contexto internacional o vivir en un país extranjero. Los alumnos adultos se enfrentan a varios retos a la hora de adquirir una segunda lengua. Algunos de ellos incluyen la adquisición de nuevos paradigmas morfológicos, estructuras sintácticas, elementos léxicos y propiedades fonológicas. A pesar de poder hablar una segunda lengua de manera fluida, es probable que el acento que retienen los alumnos adultos – con frecuencia debido a interferencia con la lengua materna (L1) – no sea tan inteligible como un acento nativo, lo que podría afectar la efectividad de la comunicación. De todos es sabido que los alumnos adultos que aprenden una segunda lengua tienen gran dificultad en dominar sonidos del habla de la segunda lengua en cuestión, y muchos de ellos retienen un acento extranjero después de haber pasado años en el país de acogida. Tener un acento extranjero puede tener repercusiones sociales. Las competencias de alumnos adultos que aprenden una segunda lengua suelen ser juzgados en base a su acento extranjero, lo que puede ser desfavorable para acceder a oportunidades profesionales, así como lograr una interacción exitosa y aceptación social.

Hace unos diez años nuevos emigrantes españoles empezaron a llegar a los Países Bajos obligados por la crisis económica de 2008 y la creciente subida de la cifra de desempleo en España. Estos emigrantes españoles están bien preparados, son flexibles, tienen altas titulaciones académicas y hablan bien inglés. La mayoría de ellos trabaja en los sectores de alta tecnología y sanidad o estudian en universidades neerlandesas. Concretamente, el 65% de los emigrantes españoles que trabajan en los Países Bajos desempeñan trabajos

profesionales o técnicos, por ejemplo, como investigadores, ingenieros, enfermeras o especialistas en tecnología informática. Sin embargo, la mayoría de ellos no tiene conocimientos lingüísticos del neerlandés cuando llegan a los Países Bajos. A pesar de poder defenderse hablando inglés, pronto se dan cuenta de la importancia de poder comunicarse en neerlandés, porque es requerido en su trabajo o porque quieren mejorar su interacción social.

Aprender neerlandés es difícil para adultos españoles – e hispanohablantes en general – y cuando se les pregunta qué les resulta más difícil a la hora de aprender neerlandés, la mayoría de ellos probablemente respondería: “la pronunciación”. El principal objetivo de este estudio es investigar los problemas de pronunciación de adultos hispanohablantes que aprenden neerlandés, y sus posibles causas, así como averiguar en qué medida los oyentes neerlandeses nativos perciben bien una pronunciación del neerlandés con acento español, en términos de inteligibilidad. Con este objetivo, se han formulado cuatro preguntas principales de investigación:

PI₁: *¿Cuáles son los errores de pronunciación más frecuentes de los adultos hispanohablantes que aprenden neerlandés, y cuáles son las causas de estos problemas de pronunciación?*

PI₂: *¿Se corresponden las vocales del neerlandés producidas por adultos hispanohablantes que aprenden neerlandés acústicamente con aquellas producidas por hablantes nativos de neerlandés?*

PI₃: *¿Son las vocales del neerlandés producidas por adultos hispanohablantes que aprenden neerlandés inteligibles para oyentes neerlandeses nativos y no expertos?*

PI₄: *¿Se corresponden las propiedades acústicas de las vocales del neerlandés producidas por adultos hispanohablantes que aprenden neerlandés con las evaluaciones perceptivas nativas de las mismas producciones vocálicas?*

El Capítulo 2 se centra en responder PI₁, o sea, *¿Cuáles son los errores de pronunciación más frecuentes de los adultos hispanohablantes que aprenden neerlandés, y cuáles son las causas de estos problemas de pronunciación?* Nuestros resultados indican que entre los adultos hispanohablantes que aprenden neerlandés, los errores vocálicos son más frecuentes, persistentes y variables que los errores consonánticos. Los alumnos hispanohablantes parecen tener problemas con contrastes relativos a la duración vocálica y al primer formante (F1), y en producir vocales anteriores labializadas, especialmente debido a la interferencia con la lengua materna. En cuanto a las consonantes del neerlandés, se han detectado problemas en fonemas individuales en posición inicial o final de palabra, pero predominantemente en combinaciones de fonemas consonánticos en posición inicial o final de palabra, a las que se debe un gran número de adiciones (p. ej., epéntesis), sustituciones y elisiones de fonemas. Como hemos mencionado anteriormente, estos errores de pronunciación son el resultado de la interferencia con el sistema fonológico de la lengua materna, y a veces del inglés, así como de restricciones en la división silábica en español (p. ej., clara tendencia a la sílaba abierta (CV) en español frente a la sílaba cerrada (CVC) en neerlandés; diferencias en la escala de sonoridad en la definición de grupos de consonantes permisibles). También se han detectado errores de pronunciación en vocales y consonantes a causa de interferencia ortográfica.

El Capítulo 3 se ocupa de la exactitud con la que alumnos hispanohablantes producen las vocales del neerlandés, ya que nuestros resultados mostraron que la mayoría de errores de pronunciación estaban relacionados con las vocales.

El objetivo del Capítulo 3 es responder PI₂: *¿Se corresponden las vocales del neerlandés producidas por adultos hispanohablantes que aprenden neerlandés acústicamente con aquellas producidas por hablantes nativos del neerlandés?* Este capítulo informa sobre una investigación en la que analizamos acústicamente las vocales del neerlandés producidas por adultos hispanohablantes y las comparamos con las producciones de hablantes neerlandeses nativos de las mismas vocales neerlandesas que teníamos como objetivo. El material acústico consistía en 29 palabras monosilábicas del neerlandés conteniendo las 15 vocales tónicas del neerlandés. Nuestros resultados han indicado que las producciones de los alumnos hispanohablantes no se correspondían a aquellas de los hablantes nativos en cuanto a valores de duración y de configuración espectral, particularmente en contrastes vocálicos que reflejan sutiles diferencias espectrales. Por ello, recurren erróneamente a la duración para realizar dichos contrastes. Por el contrario, los alumnos hispanohablantes produjeron satisfactoriamente la distinción entre vocales largas y cortas, así como las vocales largas medias y diptongos del neerlandés. Sorprendentemente, también pudieron crear una nueva categoría vocálica (una vocal anterior labializada).

Los Capítulos 4 y 5 investigan la percepción del neerlandés con acento español por oyentes neerlandeses nativos en términos de inteligibilidad, y se ocupan de responder PI₃, a saber, *¿Son las vocales del neerlandés producidas por adultos hispanohablantes que aprenden neerlandés inteligibles para oyentes neerlandeses nativos y no expertos?* En el Capítulo 4 se describe un estudio usando *crowdsourcing* en las que las mismas palabras monosilábicas producidas por alumnos hispanohablantes, y analizadas acústicamente en el Capítulo 3, eran empleadas como estímulos de habla y, acto seguido, transcritas ortográficamente por oyentes neerlandeses nativos para evaluar la inteligibilidad de las producciones de los alumnos. Una correspondencia entre

la vocal transcrita por el oyente nativo y la forma canónica (objetivo) de la misma vocal significaría que la vocal producida por el alumno hispanohablante era inteligible para oyentes neerlandeses nativos. El objetivo del estudio usando *crowdsourcing* era investigar cómo el *auris populi*, “el oído del pueblo”, evaluaría posibles producciones desviadas de vocales de una segunda lengua. Nuestros resultados indican que las vocales del neerlandés producidas por alumnos hispanohablantes fueron transcritas por los oyentes neerlandeses nativos de forma diferente a sus formas canónicas. Las transcripciones de los oyentes confirman hallazgos de investigaciones anteriores basadas en anotaciones de producciones vocálicas de hispanohablantes por expertos, a saber, que las cinco vocales del español parecen funcionar como “atractores” de la mayoría de las vocales del neerlandés. En general, nuestros hallazgos convergen con los resultados basados en datos acústicos del mismo material de habla. Un hallazgo adicional de nuestro estudio fue mostrar que la metodología *auris populi* es una valiosa herramienta para conseguir un gran número de transcripciones de habla de una segunda lengua por oyentes neerlandeses nativos y no expertos.

El Capítulo 5 investiga cómo un grupo diverso y extenso de oyentes neerlandeses nativos y no expertos percibe vocales del neerlandés producidas por alumnos hispanohablantes, una vez más siguiendo la metodología del “oído del pueblo”. En este estudio, descrito en el Capítulo 5, usamos un método empleando *snowball sampling*, que consiste en reclutar un gran número de sujetos de las redes sociales de un pequeño grupo de individuos, utilizado como punto de partida. Nuestros resultados mostraron que las vocales del neerlandés pronunciadas por alumnos hispanohablantes fueron con frecuencia transcritas de manera diferente a sus formas canónicas. Estos hallazgos consolidan resultados anteriores con respecto a la inteligibilidad del neerlandés con acento español y en la producción de habla de alumnos adultos

hispanohablantes. Los patrones de confusión vocálica observados se manifestaron con anterioridad en el estudio usando crowdsourcing, apoyando la utilidad del “oído del pueblo” para realizar futuras investigaciones focalizadas en el habla de una segunda lengua.

El Capítulo 6 intenta responder PI₄, esto es, *¿Se corresponden las propiedades acústicas de las vocales del neerlandés producidas por adultos hispanohablantes que aprenden neerlandés con las evaluaciones perceptivas nativas de las mismas producciones vocálicas?* El Capítulo 6 analiza las propiedades acústicas de las vocales del neerlandés producidas por alumnos adultos hispanohablantes e investiga cómo estas vocales son percibidas por un grupo variado y extenso de oyentes neerlandeses nativos y no expertos. Se compararon clasificaciones estadísticas de vocales obtenidas de propiedades acústicas (datos acústicos del Capítulo 3) con las clasificaciones obtenidas de los oyentes neerlandeses nativos (datos perceptuales del Capítulo 5). Asimismo, consideramos los resultados de producción y de la percepción nativa en el contexto de los niveles de dominio del neerlandés de los alumnos de acuerdo con el MCER (MCER corresponde al Marco Común Europeo de Referencia para las lenguas: aprendizaje, enseñanza, evaluación), su bagaje multilingüe, duración de residencia y uso del neerlandés. Tanto las clasificaciones vocálicas obtenidas de las propiedades acústicas de las producciones vocálicas de los alumnos (producción) como aquellas obtenidas de las clasificaciones vocálicas de los oyentes nativos (percepción nativa) fueron afectadas por el grupo de vocales incluidas como estímulo, un efecto causado por la gran variabilidad en las producciones vocálicas de los alumnos adultos hispanohablantes. A pesar de haber encontrado correspondencias entre los dos tipos de clasificaciones, también encontramos divergencias en y entre la producción y percepción dependiendo de la vocal y de los rasgos vocálicos. Acto seguido, estudiamos la variabilidad entre los alumnos hispanohablantes.

Investigamos los patrones individuales en los datos de producción y percepción nativa, y conectamos estos patrones con el nivel de dominio de neerlandés de los alumnos y con su bagaje multilingüe. Nuestros resultados con respecto a las diferencias individuales entre los alumnos hispanohablantes, tanto para los datos de producción como para los de percepción nativa, corroboraron hallazgos anteriores (cf. Burgos et al., 2014a; Flege, Frieda, & Nozawa, 1997; Munro, 1993; Yeni-Komshian, Flege, & Liu, 2000) al mostrar que la adquisición de la fonología no siempre avanza junto con el nivel de dominio de una lengua extranjera. Para concluir, el haber integrado los datos de producción y percepción nos ha proporcionado valiosos conocimientos sobre el papel de diferentes rasgos vocálicos en el aprendizaje de una segunda lengua en adultos, y de cómo estas propiedades vocálicas interactúan activamente en la manera en la que el habla de una segunda lengua es percibida. Una segunda conclusión es que mecanismos adaptativos, marcados por cambios en el límite fonémico y útiles para poder lidiar con la variabilidad de estímulos no nativos, influyen en ambos, clasificación estadística de vocales (producción) y reconocimiento humano de vocales (percepción). Por último, nuestros hallazgos indican que la variabilidad en la adquisición de la fonología de una segunda lengua es extremadamente compleja. Ocurre a diferentes niveles: en cada alumno individualmente y entre los alumnos adultos con respecto a las desviaciones segmentales según la vocal objetivo, así como en cada alumno individualmente y entre los alumnos con respecto a los rasgos (duración vocálica, uso del primer y segundo formante (F1 y F2) (por ejemplo en el caso de labilización en posición anterior) y diptongación) que los alumnos aplican para poder producir las vocales del neerlandés con exactitud.

Esta investigación ha proporcionado respuestas a las cuatro preguntas principales de investigación. Los temas más relevantes surgidos de estos

hallazgos científicos, a saber, la variabilidad en trayectos individuales de aprendizaje, mecanismos adaptativos (en clasificación estadística y reconocimiento humano), y la relación entre el nivel de dominio de una segunda lengua y el dominio de la pronunciación, serán discutidos a continuación.

Variabilidad individual

Nuestros hallazgos a la hora de explicar las diferencias individuales en el desempeño de los alumnos que adquieren la fonología de una segunda lengua han demostrado que cada uno de los 28 alumnos hispanohablantes muestra gran variabilidad en cómo producen las vocales del neerlandés. No todos los alumnos producen las mismas distinciones y no todas las distinciones están producidas con el mismo grado de distinción. Nuestros resultados evidencian que la variabilidad a la hora de adquirir los sonidos de una segunda lengua es extremadamente intrincada y se manifiesta a diferentes niveles. En concreto, existe gran variabilidad en la producción de un mismo alumno, tanto en las desviaciones segmentales (cf. Wade, Jongman, & Sereno, 2007) como en el uso de diferentes rasgos (duración vocálica, uso del primer y segundo formante (F1 y F2) (vocales anteriores labializadas) y diptongación). Este es el caso de todas las vocales objetivo. Asimismo, hay una gran variedad en la variabilidad con la que los alumnos producen con exactitud las vocales objetivo del neerlandés. La variabilidad en cada alumno individualmente y entre los alumnos lleva a diferentes patrones de confusión vocálica por vocal objetivo (cf. Bent et al., 2016; Mayr & Escudero, 2010). Los alumnos parecen recurrir a diferentes rasgos por vocal objetivo como duración vocálica, uso del primer y segundo formante (F1 y F2) (labilización en posición anterior) y diptongación, aunque el uso de estos rasgos para producir las vocales del neerlandés con exactitud no es siempre adecuado. La variabilidad también se

observa en los contrastes vocálicos que los alumnos dominan inicialmente. O sea, también existe una gran variabilidad en los diferentes trayectos que los alumnos siguen para poder pronunciar las vocales del neerlandés con exactitud.

Mecanismos adaptativos (clasificación estadística y reconocimiento humano)

La variabilidad inherente en la producción de los alumnos de una segunda lengua implica que la clasificación estadística y el reconocimiento humano tienen que adaptarse a las diferentes pronunciaciones en cada alumno individualmente y entre los alumnos para poder percibir el mensaje que se quiere emitir. Nuestros resultados muestran el funcionamiento de mecanismos adaptativos en la clasificación estadística de vocales de los datos acústicos y en el reconocimiento humano de vocales, dependiendo del grupo de vocales que se utilice (Capítulo 6). Los patrones de confusión vocálica en la clasificación estadística de vocales de los datos de hablantes nativos holandeses indica que el clasificador estadístico se adapta al tipo de datos que se incluyen en el sistema, y podría haberse adaptado o cambiado su límite de categorías fonémicas a los ambiguos sonidos de las muestras de habla no nativa. Además, la creciente mejora observada a través de las diferentes condiciones de clasificación indica que el grupo de datos acústicos que se está investigando (nativo o no nativo) puede alterar el resultado en una condición individual de clasificación (véase Capítulo 6). Nuestros resultados revelan que el clasificador estadístico es sensible al contexto en el que opera; se adapta a la naturaleza de los datos que se incluyen como *input* (datos nativos y/o no nativos). El incluir una gran cantidad de datos no nativos con gran variabilidad en los errores vocálicos parece llevar a cambios en el límite fonémico donde el clasificador estadístico acomoda datos con errores (datos no nativos) que difieren considerablemente de los datos “limpios” con sólo categorías

vocálicas objetivo (datos nativos) (cf. Berck (2017) que muestra como algoritmos de aprendizaje automático son afectados al incluir errores en datos lingüísticos).

Asimismo, en el reconocimiento humano de vocales observamos similares mecanismos adaptativos en el límite fonémico (véase Capítulos 5 y 6). La variabilidad inherente en la producción de los alumnos de una segunda lengua implica que los oyentes nativos han de adaptarse a las diferentes pronunciaciones de los alumnos. Por ejemplo, cuando los oyentes nativos escuchan habla con un acento extranjero parecen atender a detalles fonéticos que resultan de la transferencia de la lengua materna de los alumnos, para poder navegar tipos específicos de desviaciones en la señal de habla. Para poder reconocer palabras con desviaciones segmentales, los oyentes tienen que lidiar con sonidos que son versiones distorsionadas de las normas nativas, así como con sonidos que se pueden relacionar con distintas categorías fonémicas (véase Capítulos 5 y 6). En este caso, los oyentes nativos se ven obligados a alterar sus límites comunes para acomodar las ambiguas producciones no nativas que difieren de las categorías fonémicas nativas a las que están acostumbrados (cf. Bent et al., 2016; Cutler, 2012) para así poder contribuir al procesamiento lingüístico (cf. Bradlow & Bent, 2008; Cutler, 2012). Como consecuencia, los oyentes nativos tendrían que adaptarse temporalmente o cambiar su límite en la categoría fonémica para poder acomodar los ambiguos fonemas en las muestras de habla no nativa. Estos procesos podrían ayudar a comprender por qué vocales anteriores deslabializadas producidas por hablantes neerlandeses nativos fueron transcritas por oyentes neerlandeses como vocales anteriores labializadas (véase el Capítulo 5). Nuestros resultados parecen indicar lo que Cutler (2012) describe como “la plasticidad en la percepción de los oyentes adultos nativos” (p. 375). La adaptación perceptiva de oyentes nativos podría generar un

cambio en el límite entre categorías fonémicas cuando se está expuesto a habla no nativa (cf. Clarke & Luce, 2005) y nativa, especialmente cuando las vocales comparten un rasgo fonológico (cf. Chládková, Podlipský, & Chionidou, 2017), como ser vocales anteriores o posteriores labializadas. Nuestros resultados han mostrado que el oído humano nativo es capaz de procesar una gran variabilidad en la señal de habla no nativa. En conclusión, hemos observado mecanismos adaptativos en el cambio del límite de la categoría fonémica en ambos casos, clasificaciones estadísticas de vocales y reconocimiento humano de vocales.

Nivel de dominio de una segunda lengua y pronunciación

La gran variabilidad observada entre los alumnos en su desempeño de la segunda lengua, independientemente de su nivel de dominio en la segunda lengua, afecta la exactitud fonológica con la que los segmentos de la segunda lengua son producidos, y subsecuentemente, percibidos por los oyentes nativos neerlandeses. Nuestros hallazgos en cuanto a las diferencias individuales entre los 28 alumnos adultos hispanohablantes, tanto para los datos acústicos como para los perceptuales nativos, parecen indicar que la adquisición de la fonología no siempre avanza junto con el nivel de dominio de una lengua extranjera (véase Capítulo 6; véase también Capítulo 2). Tener un alto nivel de dominio en neerlandés no garantiza haber alcanzado un dominio casi nativo en la pronunciación del neerlandés (véase Capítulo 6). Algunos factores que están relacionados con nivel de dominio en una lengua extranjera son duración de residencia, y uso considerable de la segunda lengua. Estudios anteriores han demostrado que no parece ser que estos factores tengan un gran efecto en la exactitud con la que se pronuncia una segunda lengua (cf. Flege, Frieda, & Nozawa, 1997; Munro, 1993; Yeni-Komshian, Flege, & Liu, 2000).

Sorprendentemente, una descripción de los requisitos fonológicos que los alumnos necesitan cumplir para adquirir cierto nivel de dominio en la segunda lengua no están específicamente descritos en los actuales modelos de aprendizaje de segunda lengua o lengua extranjera (véase también Moyer (2013) donde se discute el tema del acento dentro de modelos de aprendizaje actuales). Esta falta de especificidad parece sugerir que la fonología no es tratada por actuales modelos de aprendizaje como un dominio lingüístico esencial, tal y como es el caso en el MCER. Los hallazgos de esta investigación han demostrado que la adquisición de la fonología de una segunda lengua es extremadamente importante para la inteligibilidad del habla de una segunda lengua y, por tanto, para poder comunicarse con éxito en el contexto de una segunda lengua. Los resultados de esta investigación ilustran igualmente la complejidad de la variabilidad individual. Es tentador concluir diciendo que los actuales modelos de aprendizaje no pueden presentar una descripción más específica del control fonológico por nivel de dominio porque la competencia oral en cuanto a la pronunciación es extremadamente difícil de definir, justamente debido a la gran variabilidad entre los alumnos de una segunda lengua.

Esperamos que los hallazgos de esta investigación contribuyan al desarrollo de herramientas de aprendizaje específicas para los hispanohablantes que deseen mejorar la exactitud de su pronunciación en neerlandés. Como ya hemos indicado, aprender neerlandés no es fácil para hispanohablantes adultos. Estamos convencidos que los resultados de esta tesis doctoral han proporcionado valiosos conocimientos sobre los problemas específicos de pronunciación que tienen los adultos hispanohablantes que aprenden neerlandés, así como de las causas de estos problemas. Estos conocimientos pueden ayudar a:

- 1) proponer direcciones pedagógicas en cuanto a la enseñanza de la fonología en las clases de neerlandés como segunda lengua;
- 2) desarrollar programas CAPT (*Computer Assisted Pronunciation Training*) específicos;
- 3) crear materiales destinados a aumentar la conciencia fonológica entre los alumnos hispanohablantes.

Esperemos que actividades de valorización ayuden a maximizar el impacto social de los hallazgos de esta investigación, y que puedan así contribuir a la integración personal y profesional de mis compañeros hispanohablantes en los Países Bajos.

Acknowledgements

Agradecimientos

Dankwoord

For your inspiration, guidance and confidence during my PhD project and on the manuscript

Roeland van Hout

For your insightful comments and suggestions during my PhD project and on the manuscript

Brigitte Planken

For serving as members of the Doctoral Thesis Committee that approved the manuscript

Mirjam Ernestus

Heike Jacobs

Hugo Quené

Juli Cebrian

Joan Carles Mora

For making my research project financially possible

Centre for Language Studies

Margot van Mulken

For your comments and suggestions at the beginning of my PhD project

Catia Cucchiarini

Helmer Strik

For helping me to complete my PhD project

George Padberg

For being my paranymphs and assisting me with my PhD defence ceremony

Núria Domínguez Rodríguez

Eric Sanders

For making me feel at home at the department of Linguistics

Nelleke Oostdijk

My colleagues

For making me feel at home at the department of Communication and Information Studies

Marianne Starren

My colleagues of the PI-group ‘Non-nativeness in communication’

My colleagues of CIW-Spanish

For their moral support, love and patience during the writing period

Sander van Loon

María Rosa Burgos Guillén

Ricardo Tejero Moya

María José Palomo Ortega

Thanks a million!

¡Un millón de gracias!

Duizendmaal dank!

Publications

- Burgos, P., Cucchiari, C., Van Hout, R., & Strik, H. (2016). Dutch vowels spoken by adult Spanish learners: A comparison between acoustic measurements and crowdsourced native transcriptions. *Proceedings of New Sounds 2016*, Aarhus, Denmark (abstract).
- Sanders, E., Burgos, P., Cucchiari, C., & Van Hout, R. (2016). *Palabras*: Crowdsourcing transcriptions of L2 speech. *Proceedings of the International Conference on Language Resources and Evaluation (LREC) 2016*, Portorož, Slovenia, 3186–3191.
- Burgos, P., Sanders, E., Cucchiari, C., Van Hout, R., & Strik, H. (2015). *Auris Populi*: Crowdsourced native transcriptions of Dutch vowels spoken by adult Spanish learners. *Proceedings of Interspeech 2015*, Dresden, Germany, 2819–2823.
- Burgos, P., Jani, M., Cucchiari, C., Van Hout, R., & Strik, H., (2014b). Dutch vowel production by Spanish learners: Duration and spectral features. *Proceedings of Interspeech 2014*, Singapore, 529–533.
- Burgos, P., Cucchiari, C., Van Hout, R., & Strik, H. (2014a). Phonology acquisition in Spanish learners of Dutch: Error patterns in pronunciation. *Language Sciences*, 41, 129–142.
- Carranza, M., Cucchiari, C., Burgos, P., & Strik, H. (2014). Non-native speech corpora for the development of computer assisted pronunciation training systems. *Proceedings of EDULEARN 2014*, Barcelona, Spain, 3624–3633.
- Burgos, P., Cucchiari, C., Van Hout, R., & Strik, H. (2013). Pronunciation errors by Spanish learners of Dutch: A data-driven study for ASR- based pronunciation training. *Proceedings of Interspeech 2013*, Lyon, France, 2385–2389.
- Burgos, P. (1995). *Los Soles Truncos: Transculturación en Puerto Rico*. *Proceedings of the UNC Romance Languages, Literature and Culture Conference*. University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States (pp. CD-rom).

Burgos, P. (1995). Apu Ollantay: Is There a Spanish Influence? *Proceedings of the Colonialism and Post-Colonialism in Latin America: Revisions and New Perspectives Conference*. Duke University. Durham, North Carolina, United States (pp. CD-rom).

Curriculum vitae

Pepi Burgos was born and raised in Spain, and has studied, lived and worked in Spain, the United Kingdom, the United States and the Netherlands. She holds a Master's Degree in Language, Literature and Culture (English and Spanish) from the University of Málaga (Spain). Pepi has worked for more than three decades as both a language and culture lecturer and a curriculum designer and developer at several universities, including Newcastle University (Newcastle, UK), Dickinson College (Carlisle, Pennsylvania, US) and the University of North Carolina at Chapel Hill (Chapel Hill, North Carolina, US). She specialises in intercultural communication, particularly – but not restricted to – issues involving Spanish-speaking communities, for example in Spain and Latin America.

At *Regina Coeli*, 'the nuns of Vught' in the Netherlands, Pepi trained international business managers on their linguistic and intercultural competences and skills. At Radboud University, she completed her PhD at the department of Linguistics, and worked as a lecturer at the department of Communication and Information Studies (International Business Communication). She now works as a lecturer and supervisor at Wageningen University & Research (the Netherlands). She teaches BA and MA courses related to Strategic and Marketing Communications, Communication and Persuasion, Research for Effective Communication, and Social Psychology. Pepi also works as an Intercultural Management Consultant for Hofstede Insights.