University of Ghent

Faculty of Arts and Philosophy

2012 - 2013

Perceiving and producing native and non-native vowels

An experimental study on the effects of first language regional variation

Master thesis submitted in partial fulfillment of the requirements for the degree of
Master of Arts in Linguistics and Literature
English - Latin

Mathijs Debaene

Supervisor: dr. Ellen Simon
Co-supervisor: prof. dr. Mieke Van Herreweghe

Acknowledgements

I would like to thank the mixture of coffee beans and nearly four decades of hip hop music (in particular the group - what's in a name - Cunninlynguists) for keeping me on track and for creating a production environment in which this endeavour could be optimally perceived.

More importantly, I would like to thank the participants from the Ghent and Antwerp area for their enthusiastic participation in the present study. I would also like to thank Katrien De Valck (Lessius University College, Antwerp) for her help in the collection of Antwerp participants and for providing an on-campus room at Lessius UC where a part of the experiments could take place.

I would like to thank my mother and father for their support, motivation and tough love during the writing process. My brother Pieterjan I would like to thank for his support and his highly necessary and greatly appreciated crash course in statistics, Ruben for his design of the time-saving Excel templates needed for the data input, and to complete the trichotomy, Arno for his support and words of nightly enlightenment.

I would also like to thank prof. dr. Mieke Van Herreweghe for the co-supervision of the present thesis. Last and foremost, I would like to express my deepest thanks to my supervisor, dr. Ellen Simon. Her speedy, yet always very thorough corrections and suggestions were highly appreciated. From the moment I was engaged to assist her with research projects as a student assistant in BA2 up until now, her enthusiasm has always been contagious, and has incited in me a passion for linguistics in general and phonology and second language acquisition in particular, which I do not doubt will reverberate throughout my life.

Table of contents

1 Introduction

The influence of regional variation in the native or first language (henceforth L1) on the target or second language (henceforth L2[1]) has long been a largely unexplored issue in the field of L2 phonology. As O'Brien & Smith remark, "[a] methodological shortcoming in previous second language (L2) acquisition studies has been that researchers have assumed an overly homogenous first language (L1) ignoring dialect differences" (2010: 297). Marinescu observes that "[i]n most second language studies, the learners' native variety is assumed to be homogenous, regardless of their regional variety" (2012: ii), and Chládková & Podlipský warn that "[n]ot taking into account learners' specific L1 dialect background could have obscured possible systematic differences between learners" (2011: 187).

Not until recently has the L1 regional variety been taken into account as a factor that influences perception (e.g. Chládková & Podlipský, 2011; Escudero, Simon & Mitterer, 2011; Escudero & Williams, 2012; Marinescu, 2012) and production (e.g. Lew, 2002; O'Brien & Smith, 2010; Marinescu, 2012) of L2 sounds. The present study is to be seen in the light of these recent developments. It builds on Debaene (2012)[2], in which the potential influence of the L1 Antwerp[3] (Brabantine) Dutch dialect on the production of the L2 English front vowels /i:/, /ɪ/ and /ɛ/ was assessed. More specifically, the study focused on whether the production of Antwerp Brabantine Dutch front vowels /i/, /ɪ/ and /ɛ/ had any influence on the speakers' production of English front vowels. The Antwerp /i-ɪ/ contrast is distinguished on a durational rather than a spectral level (while for Standard Dutch as well as English they typically are spectrally different, cf. Verhoeven & Van Bael (2002a&b) for

[1] Although the term 'second language'/L2 is used throughout this study, this term is used to refer to all language acquisition other than that of the mother tongue. English, for instance, will be the L3 for the majority of participants from this study.

[2] This is my bachelor paper ('*The effect of first language regional variation on second language vowel production. An experimental study on the production of English /i:/, /ɪ/ and /ɛ/ by Antwerp Brabantine Dutch speakers*').

[3] Note that Antwerp is used throughout this dissertation in the sense of the municipality, and does not denote the province – unless explicitly stated otherwise.

Dutch, and Collins & Mees, 2003 for English), and there is a subsequent raising of /ɛ/. The results revealed that there was some spectral overlap in the English /i:/ and /ɪ/ productions, suggesting influence of the L1 regional variety on L2 production. Results for English /ɛ/ did not reveal regional influence on production, so it was decided that the present study would focus on the Antwerp /i-ɪ/ contrast. Moreover, a second contrast is taken into consideration: according to Verhoeven & Van Bael, "[a]s far as the front and mid vowels are concerned, Antwerp [ɪ] is very close to [i] and [ʏ] is close to [y]" (2002a: 2). Antwerp realisations of close central rounded vowel /y/ and close-mid central rounded vowel /ʏ/ are produced almost identically on a spectral level, with subsequently more distinct durational features (relative to other regional Dutch varieties). Antwerp /y/ and /ʏ/ are thus contrasted in a similar way as /i/ and /ɪ/: the spectral qualities are almost identical, but /i/ is longer than /ɪ/, and /y/ is longer than /ʏ/. Neither /y/ nor /ʏ/ are part of the English vowel system, but it would nevertheless be interesting to examine what type of influence this vowel contrast has on the perception and production of English open-mid back vowel /ʌ/ and near-close near-back vowel /ʊ/. Collins & Mees remark that English learners tend to substitute English /ʌ/ with Dutch /ʏ/ (2003: 95). With respect to /ʊ/, studies have shown that there is a recent tendency in native British English to front this vowel (Hawkins & Midgley, 2005: 188), so that Dutch L2 English speakers could confuse this non-native vowel with their native /y/ sound.

The examination of potential L1 regional influence on this additional L2 vowel contrast is not the only aspect in which the present study expands upon the author's bachelor paper. An additional group of participants (with a different dialectal background) will be examined, namely people from the Ghent area. Spectral analyses from Verhoeven & Van Bael (2002a&b) show that in the East-Flemish vowel system, a qualitative distinction between /i-ɪ/ and /y-ʏ/ can be perceived (much more so than in the Antwerp dialect). Examining the L2 performances of a participant group that has L1 vowel productions with

features similar to Standard Dutch will provide us with a 'control group'. This comparative element is necessary to assess whether the results obtained in the present study can indeed be attributed to a different L1 regional background.

Another expansion is that, in comparison with Debaene (2012), the present study will also look into the issue of perception of L2 vowels. As Escudero notes, "problems producing L2 sounds originate in large measure from difficulties in perceiving such sounds in a native-like fashion" (2005: 326). In other words, L2 sound perception precedes L2 sound production. This study will thus primarily take vowel perception into account in its literature review, especially since the majority of theoretical L2 phonology frameworks are confined to L2 perception (and are only implicitly extended to L2 production). This approach is adopted by Escudero, who states that "prioritizing the role of perception in explaining the acquisition of L2 sounds is a valid and most propitious approach to the problem" (2005: 326).

Section 2 presents and contrasts three theoretical frameworks which account for L2 perception (and, sometimes implicitly, L2 production). First, Flege's Speech Learning Model (SLM, 1995) will be discussed. This model makes an explicit connection between L2 sound perception and production, which is an interesting point of view for the present study. The distinction between *identical*, *similar* and *new* L2 sounds in this model (also used in Debaene, 2012) has proven to be a methodologically useful approach. Next, Best's Perceptual Assimilation Model for non-native vowel perception (PAM), forwarded in 1995, will be discussed along with its adaptation to L2 vowel perception (PAM-L2) in Best & Tyler (2007). Thirdly, Escudero's Second Language Linguistic Perception model (L2LP, 2005) is described. Escudero's model is especially interesting, as it has already been implemented in previous studies that aim to account for the specific factor under consideration in the present study (i.e. L1 dialect and its influence on L2 sound perception), as is treated in section 3.

Section 3 presents a summary of previous findings in the domain of L2 phonological research that specifically investigates L1 dialectal differences and their influence on the perception and production of L2 sounds. Extensive attention will be paid to how theoretical frameworks are used by the authors to help elucidate their results.

Section 4 concludes the literature review of this dissertation. In this section, the relevant parts of the vowel systems of respectively Standard Dutch, Ghent (East-Flemish) Dutch, Antwerp Brabantine Dutch and Standard (RP) British English (considered to be the pre-eminent variety taught in Belgian secondary schools) are discussed. The way in which the regional Dutch varieties relate to Standard Dutch will also be discussed, as a great deal of importance is attached to the L1 production environment in the theoretical frameworks for L2 perception and production.

Section 5 presents the first of two experiments that are conducted to examine L1 Dutch regional differences in the vowel contrasts under consideration and their potential influence on L2 English vowels in two participant groups (of young Antwerp and young Ghent speakers, respectively). This section deals with perception: a Dutch as well as an English vowel categorisation task will assess how the two vowel contrasts are perceived in these two languages by the two participant groups. In the case of the Dutch /y-ʏ/ contrast, for which English has no counterpart, the perception of English /ʌ/ and /ʊ/ will be examined instead. By comparing the results of the L2 English perceptual categorisation task of the two groups, while taking into account their L1 Dutch perceptual categorisation task results, it will be examined whether L1 dialectal differences can lead to differences in the perception of L2 vowels.

In the second experiment, presented in section 6, production is examined by conducting a Dutch and an English picture-naming and sentence-reading task. This order is opted for in accordance with the factual order in which the tasks were conducted

(furthermore, L2 sound perception also precedes L2 sound production in the theoretical frameworks). In the Dutch task, productions of the /i-ɪ/ and /y-ʏ/ contrasts by the two speaker groups will be recorded and analysed through normalised spectral analysis. For the English task, the target vowels are those in the /iː-ɪ/ contrast, along with /ʊ/ and /ʌ/. Comparison of the two participant groups' productions from the Dutch task with each other and with Standard Dutch production data will reveal whether regional L1 differences can be discerned. The results of the English tasks are then compared to each other in the light of the results for the Dutch production tasks. The English results are also compared to RP English data, and discussed in terms of the concepts and theoretical models discussed in the literature review.

2 Theoretical models of L2 perception and production

2.1 Flege's Speech Learning Model (SLM, 1995)

2.1.1 Context of the model's origin

One of the most frequently cited L2 perception and production models is Flege's Speech Learning Model (SLM, 1995). An important assumption made by Flege is that

> *"the phonetic systems used in the production and perception of vowels and consonants remain adaptive over the life span, and that phonetic systems reorganize in response to sounds encountered in an L2 through the addition of new phonetic categories, or through the modification of old ones"* (1995: 233).

This assumption is a reaction against the Critical Period Hypothesis, which posited that for L1 and L2 acquisition there exists a window of opportunity (a critical period) regarding perfect acquisition. When this window – which is closely related to neurological maturation (Flege, 1995: 234) and thus to the age of learning - closes, ultimate attainment in L1 and/or L2 becomes impossible. In Flege's words, "[m]any believe that new forms of speech cannot be learned perfectly once a critical period has passed" (1995: 234). Although Flege (1995) does not argue against the scientific, neurological basis of these theories and hypotheses, he does argue against an absolute validity of these assumptions. Empirical evidence in the field of L2 acquisition supports his claims. First of all, he presents results from a study conducted by Flege, Munro & Mackay (1995), where he used ratings of the average degree of perceived foreign accent in the English of Native Italians who migrated to Canada. The results of this study suggest that there is by no means an abrupt break (implying a critical period) regarding the degree of perceived foreign accent (and thus the ability of the L2 learners to acquire and produce an L2). On the contrary, a linear relation between the age of arrival in Canada and the average degree of perceived (foreign) accent becomes visible in the analysis of the results. A second argument against the Critical Period

Hypothesis lies in the observations Flege gathered from other studies: Flege states that "[a]lthough very common, foreign accents are apparently not *inevitable*" (1995: 236). It is thus possible for an adult L2 learner to attain accent-less proficiency of a second language.

Flege furthermore points out that in previous views of L2 acquisition, age of learning, L2 experience and other external factors were commonly ignored in the description of L2 production errors (1995: 235). This rather static view of foreign accent (which Flege encounters in, among others, Weinreich, 1953 and Lehiste, 1988) proclaimed that the principal phonological cause of foreign accent was L1 interference, with theories based solely on this particular element of L1 interference. These theories are criticised by Flege, as they do, for instance, not take into account factors such as the age of learning, how long the L2 has been spoken, or which people it was spoken with (1995: 235). With this criticism in mind, the questionnaire (regarding L1 and L2 background) contained in the experiment (cf. Appendix A) is considered essential to account for possible individual variation in the experiment results.

At the basis of Flege's Speech Learning Model lies a paradox: "At an age when children's sensorimotor abilities are generally improving, they seem to lose the ability to learn the vowels and consonants of an L2" (1995: 234-35). In this remark, Flege criticises the view that one of the requirements for the accurate production of an L2 sound would be "the learning of gestures with which to reliably reproduce the represented L2 sounds" (1995: 236), which runs counter to Best's PAM (1995) (cf. Section 2.2 below). Flege argues that, although foreign accents and misarticulations are related to motoric difficulties, these motoric difficulties are not the inherent cause of foreign accent (1995: 236). Instead, Flege (1995) searches for the roots of foreign accent and misarticulations in the perception of a second language.

In this model, the presumed cause for (many) of the L2 production errors is a failure in the correct perception of L2 sounds. Since both production and perception are considered in the present study, this connection between L2 perception and production is highly relevant. On the one hand, Flege does not object to Locke's proposition that "the root cause of many L1 segmental production errors is to be found at a "motor" level rather than at a "mentalistic (level) of linguistic organization [i.e. thus a perceptual basis]" (1980: 465, as cited in Flege, 1995), he also ascertains that this conclusion cannot simply be extended to L2 learning. Flege notes that "bilinguals tend to interpret sounds encountered in an L2 through the grid of their L1 phonology" (1995: 237). This consideration anticipates the substantial role of L2 sound perception in Flege's model (1995), the extent of which will accurately be assessed in section 2.1.2.

2.1.2 SLM postulates and hypotheses

2.1.2.1 SLM postulates

The SLM, as proposed in Flege (1995), advances four major theoretical postulates to account for L2 sound acquisition, from which, as will be discussed in section 2.1.2.2, seven hypotheses arise (six of which pertain to perceptual characteristics of L2 acquisition). Together, they constitute the SLM as a theoretical framework.

One of the most significant elements to be derived from the postulates are that L2 learning takes place via a process similar to L1 learning (implying among other things a mutual influence of L1 and L2 on each other), and that contrast between the L1 and the L2(s) is to be maintained. Another element of major importance is the concept of *phonetic categories* that is advanced. These phonetic categories are "long-term memory representations" in which "[l]anguage-specific aspects of speech sounds" are postulated to be specified (Flege, 1995: 239). For Flege, the SLM (1995) takes place on the phonetic rather than the phonemic level.

For an overview of the four postulates we refer to Flege (1995: 239).

2.1.2.2 SLM hypotheses

Backed by empirical evidence and with the postulates as foundation, Flege forwards seven hypotheses as part of the SLM (1995), six of which predict how L2 sound perception emerges and develops itself. It should always be kept in mind that these hypotheses apply to the perception and production in experienced learners. Flege (2005a), for instance, admits that in early stages of L2 learning L2 sounds are usually perceived as instances of a previously established L1 category, even if there are phonetic dissimilarities between an L1 and an L2 sound. However, as the participants in the present study are considered experienced L2 English learners, the hypotheses hold for them.

Due to limited space and the fact that Flege's assessment of the hypotheses is already very concise (yet thorough), we refer to Flege (1995: 237 and further) for an overview and analysis of the SLM hypotheses. Important is that the level of abstraction in Flege's SLM is "a position-sensitive allophonic level, rather than at a more abstract phonemic level" (1995: 239). A new phonetic category is constructed if "bilinguals discern at least some of the phonetic differences between the L1 and L2 sounds" (Flege, 1995: 239), and "the greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned" (Flege, 1995: 239) and thus that a new phonetic category is constructed for the L2 sound. Age of learning (AOL) also plays a role, in that "[t]he likelihood of phonetic differences between L1 and L2 sounds, and between L2 sounds that are non-contrastive in the L1, being discerned decreases as AOL increases" (Flege, 1995: 239). It should be noted that the formulation of this hypothesis leaves open the possibility for someone with an adult AOL to attain an accent-less L2 proficiency. This would leave open the possibility for Antwerp and Ghent dialect speakers to efficiently perceive and produce ways of contrasting between L2 vowels that are not present in their L1 dialect.

Another important SLM concept is equivalence classification, a mechanism through which, as Flege states, "[c]ategory formation for an L2 sound may be blocked" (1995: 239). "When this happens," Flege posits, "a single phonetic category will be used to process perceptually linked L1 and L2 sounds (diaphones). Eventually, the diaphones will resemble one another in production" (1995: 239). Due to the diaphones being perceptually linked, Flege furthermore leaves open the possibility of a mutual influence of L1 on L2 as well as of L2 on L1.

Attention is given to the cause of foreign accent as well. Flege states that

> *"[t]he phonetic category established for L2 sounds by a bilingual may differ from a monolingual's if: 1) the bilingual's category is "deflected" away from an L1 category to maintain phonetic contrast between categories in a common L1-L2 phonological space; or 2) the bilingual's representation is based on different features, or feature weights, than a monolingual's"* (1995: 239).

The first possibility again stresses the mutual relationship between L1 and L2 sounds. This deflection could subsequently cause the established L2 phonetic category to differ from that of a monolingual speaker, as this monolingual speaker does not experience any interference between different cross-linguistic phonetic categories in the same phonological space. Because of the emphasis on interaction between L1 and L2, it is implied that L2 speakers with different L1 backgrounds perceive and produce the same L2 sound differently. It also (implicitly) assigns an important role to the L1 background: emphasis on different L2 features or feature weights could arise based on which features or feature weights are important in the L1. The durational distinction rather than the spectral distinction between the Antwerp /i-ɪ/ and /y-ʏ/ vowel pairs could prove to be an example of such a presentation based on different feature weights. The weight of the durational feature could be bigger than the weight of the spectral feature in these two vowel pairs,

which are primarily contrasted on a durational rather than on a spectral level (as opposed to, for instance, Standard and Ghent (East-Flemish) Dutch).

The SLM framework, which is rooted in L2 perception, is also extended to the production of L2 sounds. This L2 production very straightforwardly derives from L2 perception, in that "[t]he production of a sound eventually corresponds to the properties represented in its phonetic category representation" (Flege, 1995: 239). This explicit link between perception and production of L2 sounds suggests the fruitfulness of the present study with its focus on both perception and production.

2.1.3 Identical vs. similar vs. new sounds

From the SLM as it is presented in Flege (1995) follows that, with the L1 sound inventory as reference, there are three possible types of phonetic relationship between an L2 and an L1 sound.

An L2 sound may be *identical* to an L1 sound, meaning that there is "no difference, or [the difference is] too small to detect auditorily" (Flege, 2005a). *Identical* L2 sounds are transcribed with "the same IPA symbol used to represent a sound in the L1" (Flege, 1997: 17), and they are believed to be perceived and produced accurately by L1 learners, "as the result of a process referred to as "positive transfer""(Weinreich, 1953, as cited in Flege, 1997: 17).

An L2 sound can also be *similar* to an L1 sound: it is "represented by the same IPA symbol as the L1 sound, even though statistical analyses reveal significant – and audible – differences between the two" (Flege, 1997: 17). The amount of L1-L2 phonetic difference is larger. In this case, the phonetic category established for the L1 sound and the L2 sound residing in the same phonological space do not entirely overlap (or at least not to as absolute an extent as is the case for *identical* sounds).

Thirdly, an L2 sound can be entirely *new* with regard to the L1 sound inventory. A *new* L2 sound, which evidently differs depending on the native language of the L2 speaker, "has no obvious phonetic "counterpart" in L1" (Flege, 2005a), and it "differs acoustically and perceptually from the sound(s) in L1 that most closely resemble(s) it" (Flege, 1997:17). However, "unlike a *similar* sound, it is represented by an IPA symbol that is not used for any L1 sound" (Flege, 1997:17).

An implication of the fact that age of L2 learning plays a role is that "phonetic categories needed to produce and perceive L2 sounds can be added readily until the age of 5-6 years, when the phonetic system begins to stabilize" (Flege, 1997: 12). "After this age," Flege continues, "supplementary phonetic categories can be established for the types of sounds categorized as *new*, but not for the types of sounds categorized as *similar"* (1997: 12). This is due to the mechanism of equivalence classification. Flege points out that this "basic cognitive mechanism" also allows speakers, for instance, to "use the word *chair* correctly in identifying the many physical exemplars of this furniture type" (1997: 13). Accordingly, speakers can group the varying acoustic instantiations of an L1 sound into the relevant phonemic categories. However, this could prove to be an obstacle for L2 learning, especially after the 'critical period' of 5-6 years, and *similar* L2 sounds are equivalently classified as the L1 sound without more detailed refining.

As the participants of the present study almost exclusively started their (formal) English education around the age of 12 (in the first or second year of secondary school - cf. Appendix A), equivalence classification is not to be excluded in the present study[4]. In the application of Flege's SLM (1995) to the results of the present study, it will therefore be essential to determine if the cross-dialectal differences between the L1 Dutch speakers

[4] However, Flemish children are likely to be exposed to English before the age of twelve, due to, for instance, media and pop culture.

result in a different view of the English vowels under consideration as *similar*, *new* or *identical*.

It should be noted that the IPA criterion for distinguishing the three different categories of L2 sounds is not as proper a criterion as, for example, spectral and durational analysis of sounds. Furthermore, it should be noted that the three-way distinction between the different categories is merely a methodological simplification: L1–L2 phonetic dissimilarity should be regarded "as a continuum, not as tripartite *identical-similar-new* division" (Flege, 2005b). For methodology's sake, this three-way distinction will be maintained when making predictions for the L1 Dutch speakers under consideration in the present study. Wherever this is possible, degrees of similarity of the different dialectal L1 vowels with respect to the English L2 vowels will be used as a reminder of the continuous character of the different categories in which L2 sounds can be mapped phonetically. This way, an L1 vowel of the Ghent (East-Flemish) Dutch dialect, for instance, may or may not be 'more *similar*' to an L2 English vowel than its counterpart from L1 Antwerp Brabantine Dutch.

2.1.4 Remark on the SLM with regard to the present study

As stated in the introduction (section 1), one of the shortcomings of Flege's SLM (1995) is that it does not mention the different L1 dialectal backgrounds of L2 speakers. However, just like the sharp distinction between *new*, *similar* and *identical* sounds, this should be viewed as a methodological simplification rather than an inherent shortcoming in the theory. If research proves an L1 dialect (rather than one standard L1 variety) to be predominant in the speaker's daily language use, this different L1 variety may indeed be viewed as the speaker's first language, on the basis of which the L2 phonetic categories are constructed, and by which these L2 phonetic categories can be influenced accordingly.

2.2 (L2) Perceptual Assimilation Model (PAM, 1995/2007)

2.2.1 Introductory remarks

The Perceptual Assimilation Model, as it was devised by Best in 1995, is a theoretical model of non-native speech perception which provides us with a well-structured examination of the possible scenarios in encountering non-native sounds. Despite its seeming suitability as a framework to account for the results of the present study, the 1995 version of the model focuses solely on non-native speech perception of inexperienced, naïve listeners. Language users from this target group are defined as "functional monolinguals, i.e., not actively learning or using an L2, and are linguistically naïve to the target language of the test stimuli" (Best & Tyler, 2007: 6). As the participants for the present study all received at least 5 years of formal English education at secondary school (cf. Appendix A), these L2 speakers cannot be categorised as naïve listeners. Best's PAM (1995), however, has been revised in 2007, in an attempt to extend its range to L2 learners, where it was found that the scenarios constructed for naïve listeners can also be applied to more experienced L2 learners. This Perceptual Assimilation Model for L2 Learners (PAM-L2, Best & Tyler, 2007) furthermore takes into account the postulates and predictions of Flege's SLM (1995) and assesses them from a PAM (Best, 1995) perspective, making PAM-L2 especially relevant for the present study. In the present discussion, the framework of the 1995 version of Best's PAM will therefore be treated first; subsequently, Best & Tyler's Perceptual Assimilation Model for L2 learners (PAM-L2, 2007) and how it incorporates nonnative perception by L2 learners will be discussed.

2.2.2 Best's Perceptual Assimilation Model (1995)

2.2.2.1 Key points

Best's Perceptual Assimilation Model (1995) operates on a gestural phonetic level. According to Fowler & Rosenblum, phonetic gestures are "organized movements of one or

more vocal-tract structures that realize phonetic dimensions of an utterance" (1989: 102). The model thus, in conformity with Flege's SLM (1995), puts emphasis on the phonetic level of analysis. Unlike Flege, however, Best does not exclude from her framework the phonological level of analysis: she asserts that there is "a common gestural domain for both phonetic details and phonological structure, in which the constellations of language-specific gestural details *are* the phonological elements of the language" (1995: 183). This is particularly relevant in the present study, as the Antwerp (Brabantine) Dutch dialect distinguishes some of its vowel contrasts mainly with the gestural feature of vowel duration, unlike the equivalent Ghent (East-Flemish) Dutch dialect vowel pairs, which would be contrasted primarily on a spectral rather than a durational level.

As in Flege's 1995 SLM, Best asserts that in PAM (1995), the perceived non-native sounds are filtered through the framework of the native language. She points out that

> *"[t]he fundamental premise of the perceptual assimilation model of cross-language speech perception is that non-native segments [...] tend to be perceived according to their similarities to, and discrepancies from, the native segmental constellations [i.e. the L1 phones] that are in closest proximity to them in native phonological space"*(1995: 193).

In this sense, the non-native sounds (with other segmental constellations and thus phonetic realisations) are perceptually assimilated to native phonological (not phonetic) categories.

2.2.2.2 Three assimilation patterns

Best discerns three different patterns for the perceptual assimilation of non-native segments. First of all, the non-native segment or phone can be "assimilated to a native category" (Best, 1995: 194). In this particular case, the non-native segment may be heard as a "good exemplar of that category", "an acceptable but not ideal exemplar of the category", or "a notably deviant exemplar of the category".

Second, the non-native segment may be "assimilated as uncategorizable speech sound" (1995: 194). In this case, the non-native segment is "assimilated within native phonological space as a speechlike gestural constellation, but not as a clear exemplar of any particular native category" (1995: 194). The NL segment is thus still recognized as a speech sound when processed through the native phonological space, yet it "falls [...] in between specific native categories" (1995: 194).

In the third scenario, this recognition as a human speech sound is not present: the non-native segment is "not assimilated to speech", but is "heard, instead, as some sort of nonspeech sound", when projected upon the phonological space of the native language (1995: 195).

2.2.2.3 Types of assimilation of non-native contrasts

As the present study assesses the perception of several vowel contrasts, it is interesting to note that Best's PAM (1995) also touches upon the assimilation of non-native contrasts: Best states that "[a]ssimilation patterns for non-native *contrasts* follow predictably from the assimilation of each member of the contrast" (1995: 195). She asserts that PAM can predict "the degree of perceptual differentiation, or discriminability, for diverse non-native contrasts" from "the assimilation of each of the contrasting non-native segments" (1995: 195). She then provides a concise overview of the different patterns of pairwise assimilation. The perception experiment in the present study is designed in such a way that the participants are meant to assume that the perceived vowels are indeed human speech sounds, which they furthermore are forced to categorise as an L1 vowel. Therefore, this summary will leave out the scenarios in which at least one of the segments in the non-native contrast is assimilated as an *uncategorizable* or as a *nonspeech sound*.

In the Two-Category assimilation (TC Type), "[e]ach non-native segment is assimilated to a different native category, and discrimination is expected to be excellent"

(Best, 1995: 195). If there is a *Category-Goodness Difference (CG Type)*, the two sounds of the contrasts are "assimilated to the same native category, but they differ in discrepancy from native "ideal"" (Best, 1995: 195). In other words, one of the non-native speech sounds of the contrast will be considered by the listener as a *good* example of the phone which constitutes the phonological space to which both sounds are assimilated, while the other non-native speech sound of the contrast will be regarded as *deviating* from the phone in this native language phonological space, or at least as a less good example of it. Discrimination is expected to range from "moderate to very good", and depends on the "magnitude of difference in category goodness for each of the non-native sounds" (1995: 195). The perception experiment in the present study, however, was not designed to assess 'goodness' judgments, so it will not be possible to provide evidence for this type of assimilation on the basis of the experiments in the present study.

The third type of assimilation relevant to the present study is the "*Single-Category Assimilation (SC Type)*", which corresponds to the CG type of assimilation in that both non-native sounds are assimilated to the same native phonological category. SC assimilation differs, however, from CG assimilation in that both sounds of the contrast are perceived as "equally discrepant from the native "ideal"; that is, both are equally acceptable or both equally deviant" (Best, 1995:195). Best concludes that a poor discrimination is expected (1995:195).

2.2.3 Best & Tyler's PAM-L2 (2007)

2.2.3.1 Background

As mentioned before, PAM as it was devised by Best in 1995 does not take into account the perception of second language sounds by more experienced L2 learners (the group on which Flege's 1995 SLM focuses). It should be noted that the group of L2 speakers and listeners in the present study cannot be classified as 'experienced' in the sense in which it is

used by Best & Tyler in their PAM-L2: a minimum of six to twelve months of L2 immersion is required (2007: 14). Apart from the one participant who spent three months in England, none of the participants spent such an amount of time in an English-speaking country. However, the participants did receive at least five years of formal English education at secondary school level. (cf. Appendix A). Though a simplification, it is nevertheless safe to assume that when original PAM can be extended to experienced L2 learners, it can also be relevant to account for the results of the L2 listeners in the present study (who cannot be classified as 'experienced' L2 learners, but can neither be seen as the functional monolinguals for which Best's PAM (1995) was devised).

Best & Tyler point out that non-native and L2 speech perception "have frequently been assumed to be essentially identical, that is, to reflect the same L1 influences" (2007: 15). They continue with the observation that because of this assumption, the two models that are most frequently cited in this way, namely Flege's SLM (1995) and Best's Perceptual Assimilation Model (1995), have been wrongly used interchangeably, as if these models both address the two situations at hand (2007: 15). Best & Tyler do not argue that this assumption is completely incorrect, yet they state that the connection has not been assessed before. For this reason PAM-L2 was devised, as its aim is "to probe the commonalities and complementarities of the two models, and especially to explore whether and how SLM can be used as a starting point to extend PAM's nonnative speech perception framework to L2 learners" (2007: 15).

2.2.3.2 PAM-L2: the SLM postulates from a PAM angle

Best & Tyler assert that "neither PAM nor SLM restrict their predictions about native language influences solely to the influence of phonological contrasts in the L1" (2007: 15). Section 2.1 reveals that Flege's SLM (1995) operates mainly on a phonetic level. Best's PAM (1995), on the other hand, "has characterized the nature of L1 influence as being based on perceptual learning of phonetic-articulatory patterning at both the abstract contrastive [i.e.

phonological] level and, importantly, at the level of non-contrastive gradient phonetic detail" (Best & Tyler, 2007: 16). In other words, PAM views the phonetic and phonological level as intertwined. With this observation in mind, Best & Tyler (2007) critically assess the SLM postulates as presented in Flege (1995).

Best & Tyler generally agree with Flege's first SLM (1995) postulate regarding the lasting intactness of L1 sound system learning mechanisms and processes, and their applicability to L2 learning. They point out that during their life span, speakers can adapt fairly easily to "changes in the ambient language environment" like L1 dialectal innovation or variation, by "perceptual learning of new higher-order invariants [i.e. phones] over age" (Best & Tyler, 2007: 19). L2 learning is then seen as "a functional extension of [this process]" (Best & Tyler, 2007: 19).

On a theoretical level, PAM is based on a direct realist view of sound perception (Best, 1995). In short, this means that the perceived sounds are viewed and treated by Best & Tyler (2007) as the actual sounds themselves – not as mental representations of these sounds by the perceiver. Yet in the second postulate of the SLM, the concept of phonetic categories as "long-term memory representations" in which "[l]anguage-specific aspects of speech sounds are specified" is introduced (Flege, 1995). This "assumption of mental representation" is thus in conflict with PAM (Best & Tyler, 2007: 20). In PAM-L2, there is simply no need for these mental representations in phonetic categories: "Rather, the listener directly perceives the articulatory gestures of the speaker and, through perceptual learning, comes to detect higher-order articulatory invariants in speech stimuli" (Best & Tyler, 2007: 20). Once again, an interaction between the different levels of analysis is assumed here: "Language-relevant speech properties are differentiated not only at the phonetic level but also at the higher-order phonological level, as well as at the lower-order gestural level" (Best & Tyler, 2007: 20). The ability of L2 learners to correctly discriminate L2 sounds is influenced by each of these three levels, in a separate way or combined. Best &

Tyler underline the importance of these "levels of attentional focus, and the relationship between the L1 and the L2 at each": they are "essential to explaining L1 and L2 perceptual learning from common theoretical principles" (2007: 21).

The third SLM postulate builds on the second one, asserting that L1 phonetic categories evolve over the life span, and are adaptable so that "the properties of all L1 or L2 phones identified as a realization of each category" (Flege, 1995: 239) are reflected. Because this postulate expands on the concept of phonetic categories, it is incompatible with Best's PAM (1995). However, the idea behind this postulate remains intact when reformulated using PAM-L2 terms: "exposure to L2 phones that are assimilated to an L1 phonological category may require the perceiver to discover a different set of invariants to encompass the new shared phonological category" (2007: 21). In the treatment of this postulate, Best & Tyler furthermore explore the (previously largely disregarded) issue of "how listeners identify nonnative phones as equivalent to L1 phones, and the level(s) at which this occurs" (2007: 21-22). They conclude, again, that phonetics as well as phonology play an important role:

> *"contrasts at the functional linguistic level of the L1 phonology and their relationship to phonological contrasts in the L2 are as important to perceptual learning as phonetic categories [note that Flege's sense of mental representation is not implied here] in the two languages according to PAM but not SLM"* (2007: 22).

The fourth postulate of the SLM states that bilingual speakers try to sustain a contrast between the phonetic categories of a first and a second language that exist in a common phonological space (Flege, 1995: 239). The central point of critique of Best & Tyler (2007), that Flege's SLM (1995) pays little to no attention to the phonological level of analysis in its explanations, recurs again. Best & Tyler essentially agree with this postulate, yet they point out that "both phonetic *and* phonological levels interact in L2 speech

learning, and that this depends crucially on the relationship between the phonological spaces of the L1 and L2" (2007: 22). Best & Tyler present a theoretical scenario. If perceptual assimilation takes place of a phonological category of the L2 to an L1 phoneme, and if the respective phonetic realisations of this shared category are discriminable, they posit that "the listener should be able to maintain the L1 and L2 phones as separate phonetic realizations of the one phonological category" (2007: 23). Another interesting aspect of this scenario is that it implies a possible influence not only of the L1 on the L2, but also of the L2 on the L1: "the listener could become increasingly perceptually attuned to this phonetic distinction within the single phonological category" (Best & Tyler, 2007: 23). This could subsequently cause a shift "in the exact details of the L1 and L2 phoneme" (Best & Tyler, 2007: 23) when compared to people who lack this type of attunement, namely the respective monolingual speakers of the L1 and L2 under consideration. In principle, this implies that learners of different L2's could perceive the same L1 differently. Best & Tyler conclude that "perceptual changes may arise as a result of having two phonetic categories for a single interlanguage phonological category" (2007: 23).

2.2.3.3 L1-L2 scenarios in PAM-L2

Best & Tyler (2007) move on to present a systematic assessment of the possible scenarios of how L1 and L2 sounds relate at the phonological level. From this perspective, they predict "the likelihood of success of perceptually learning an L2 phonetic category [in Best's 1995 PAM sense of 'set of gestural features attuned to be distinguished increasingly well, but not stored as memory representations, by the perceiver'] in each case" (Best & Tyler, 2007: 24).

The main issue is then "whether or not the learner has perceived equivalence between an L2 and an L1 phonological category, that is, has perceptually assimilated the L2 phone to this L1 phonological entity" (Best & Tyler, 2007: 24). This process would then be similar to Flege's equivalence classification (1995), but with both the phonological and phonetic planes as relevant levels of analysis. This implies that an L2 phone could be

assimilated to an L1 phone if the "phonological category has a similar contrastive relationship to surrounding categories in the phonological space" (Best & Tyler, 2007: 24). However, as opposed to what Flege's SLM (1995) would suggest, this need not imply that there should also be phonetic similarity between the L1 and L2 phones. Best & Tyler present the example of French uvular /r/: although this phone is (as a rule) perceptually assimilated to English liquid /r/, the two phones do not resemble each other on a phonetic level (2007: 24-25). A similar contrast was impressionistically heard in the two groups of participants in the present study, too. The majority of Ghent speakers produce the uvular variant of the rhotic consonant /r/ (viz. [ʁ]), while all of the Antwerp speakers use an alveolar trill [r] in the Dutch production ta sk. Both participant groups nevertheless accurately produced the English alveolar approximant [ɹ]. Since it is generally assumed that perception precedes production (e.g. Escudero, 2005: 326), the Ghent production evidence can be regarded as an extension of the perceptual assimilation of the Ghent rhotic consonant and the English equivalent on a phonological level.[5]

Just like Best's original Perceptual Assimilation Model for naive listeners (1995), Best & Tyler's PAM-L2 (2007) works with a framework in which the different cross-language scenarios regarding L2 are assessed by means of looking at how L2 minimal contrasts are perceived by L2 learners. These scenarios are extensions of possible 1995 PAM non-native speech perception scenarios to L2 learning. The first scenario is one in which "*[o]nly one L2 phonological category is perceived as equivalent (perceptually assimilated) to a given L1 phonological category*" (Best & Tyler, 2007: 25). Fine-grained perceptual learning of the contrast component which is perceptually assimilated to the L1 category is not expected to take place. There is no need to do so: the other part of the minimal L2 contrast is not perceptually assimilated to the L1 phonological category at hand. It is expected that L1-to-

[5] It should be noted that orthography also plays a role: the participants are undoubtedly aware of the fact that English (like Dutch) also uses <r> for its rhotic consonant; their English productions are thus not only based on perception.

L2 perceptual assimilation will take place both on the phonological and the phonetic level (Best & Tyler, 2007: 26). Another possibility is that the L2 sound is only assimilated on a phonological level (cf. the discussion of rhotic consonants in the previous paragraph).

In the second case, "*[b]oth L2 phonological categories are perceived as equivalent to the same L1 phonological category, but one is perceived as being more deviant than the other*" (2007: 26). In this case, a "category goodness assimilation contrast" occurs: one of the parts of the L2 minimal contrast is seen as a better representative of the L1 phonological category than the other part. Best & Tyler predict that the L2 phones will be discriminated fairly well, "though not as well as two category assimilation types" (2007: 27), and that the lexical and functional differences between the two L2 phones can be discerned fairly easily by listeners. New phonological and phonetic categories are likely to be formed for the L2 phone that deviates most from the L1 category. A new phonological category is not formed for the L2 phone of the contrast that is closest to the L1 category, but a new phonetic category can still be formed depending on the "degree of its perceived similarity to the L1 category" (Best & Tyler, 2007: 27). The more deviating from the L1 phonetic category, the more likely a new phonetic category will be formed for the L2 phone.

The third situation presents a scenario where "*both L2 phonological categories are perceived as equivalent to the same L1 phonological category, but as equally good or poor instances of that category*" (Best & Tyler, 2007: 27). This is a case of "single-category L2 contrast assimilation", where the two phones of an L2 minimal contrast are perceptually assimilated to the same L1 phonological and phonetic category (Best & Tyler, 2007: 27). Best & Tyler expect rather poor performance of L2 learners when perceiving single-category assimilated phones, and they expect that two words which only differ in the minimal contrast concerned will be perceived as homophones (2007: 27). They point out that a lot depends on lexical factors: the more this perceived homophony hinders L2 communication, the greater

the need will be for L2 learners to perceptually differentiate between the two minimally contrasted phones.

Finally, Best & Tyler describe the case where there is no L1-L2 phonological assimilation, which occurs "if the naive listener does not perceive either of the contrasting L2 phones as belonging clearly to any single L1 phonological category, but rather as each having a mixture of more modest similarities to several L1 phonological categories" (2007: 28). The phones, in this scenario, are uncategorised, and "one or two phonological categories may be relatively easy to learn perceptually" (2007: 28). However, the closer these uncategorised L2 phones are to each other in the phonological space, the harder they will be to perceptually discern. Best & Tyler furthermore point out that these types of L2 phones (which would be called *new* in Flege's 1995 SLM paradigm) do not, as Flege (1995) suggests, bear upon resemblances or differences with the *closest* L1 phonetic category : the whole of "comparative relationships within the interlanguage phonological system" is essential here (2007: 28).

In conclusion, Best & Tyler's PAM-L2 (2007) offers a suitable theoretical and methodological framework to help elucidate the results of the present study. It primarily differs from Flege's SLM (1995) in the sense that this model pays attention to the phonological level of analysis as well. The preceding discussion demonstrated that the two models are fairly compatible with each other. In the case study it will be fruitful to aim to explain the results of the perception tasks in the light of both these models.

2.2.3.4 Considerations with regard to the present study

A distinction is made by Best & Tyler (2007) between phonological and phonetic categories (not to be confused with Flege's phonetic categories) of speech information. This division follows the traditional sense in which the levels of phonology and phonetics are differentiated. Phonological categories pertain to "speech information that is relevant to

minimal lexical differences in a given language", while phonetic categories "do not signal lexical distinctions, but may instead provide perceptual information about the speaker's identity, or their region or language of origin" (Best & Tyler, 2007: 21).

Henceforth, the level of analysis in the present study on which most dialect-influenced differences take place is expected to be the phonetic one. Although Best & Tyler's PAM-L2 (2007) does not touch upon production, L1 and L2 sounds still need to be produced in order for these sounds to be perceived. Therefore, it seems permitted to extend the consideration at hand to the level of production: special attention should be paid to the L1 and L2 production tasks, as different (phonetic) sets of invariants (namely vowel quality and vowel duration) are expected to play an essential role in the differentiation of speaker productions from the two regional participant groups.

Concerning the perception tasks, more caution is needed. The tasks work with nonsense words, so that no clear-cut lexical information (that could facilitate the process of correct vowel categorisation) can be derived from the vowel context. In the perception experiment, more attention is to be paid to the relation between the respective L1 dialects on the one hand, and L2 English on the other hand. L1 dialectal differences could be substantial enough to give rise to different preferences of the two groups of participants in selecting the appropriate phonological categories available in the tasks. This, in turn, could imply a different perceptual assimilation pattern between the two participant groups.

In section 2.3, a third and final L2 speech perception model is critically assessed, namely the Second Language Linguistic Perception Model (L2LP) by Escudero (2005).

2.3 Escudero's Second Language Linguistic Perception Model (L2LP, 2005)

2.3.1 Escudero's (2005) view of language perception

Escudero's doctoral dissertation (2005) advanced the Second Language Linguistic Perception Model (L2LP), the most elaborate theoretical L2 perception framework so far. The name of this model reveals that L2LP is situated on a phonological (i.e. linguistic) level of analysis (unlike, for instance, Flege's 1995 SLM, which is situated on a phonetic level). Before L2LP is described, Escudero (2005) advances and incorporates several theoretical concepts essential to the foundation of L2LP. These concepts are initially applied to L1 perception, but are subsequently extended to L2 perception.

If sound perception is linguistic (or phonological) instead of phonetic, this implies that speakers use what Escudero calls a language-dependent perception grammar, which "performs the mapping of the [speech] signal through constraints that map or connect the acoustic properties of the input with sound representations" (2005: 44), and which are separate for each language (2005:118). These constraints determine the boundaries between which a sound in the (phonetic) speech signal can be situated in order to still be regarded as an acceptable representative of the (phonological) sound representation associated with the space within these boundaries. Such constraints lie along different auditory 'cues' which can be spectral in character, e.g. vowel height (F1) or backness (F2), but also durational (Escudero, 2005: 21).

Users of different languages moreover attach importance to different cues. For instance, the perception grammars of L1 speakers of languages with only short vowels are not expected to focus on constraints which attend to durational vowel features. The durational cue constraint will then have to be acquired before accurately perceiving vowels in a language with durational distinctions.

To further theoretically clarify how sounds are successfully perceived, Escudero integrates the Optimal Perception Hypothesis in L2LP (2005). The way in which an auditory speech signal is mapped onto the different phonological categories of the listener, depends on "the specific characteristics of the listener's production environment" (Escudero, 2005: 52). In other words, the assertion is again made that perception is based on language-specific aspects. Escudero points out that, as a way for listeners to optimally perceive a speech signal, they always seek to "perceive the sounds of a language [...] by making categorization decisions that lead to *maximum-likelihood behaviour* [...] which minimizes the possibilities of misunderstanding a speaker" (2005: 52). This Optimal Perception Hypothesis, which she derives from Boersma (1998), posits that "experience with the particular [i.e. language-specific] ways in which sounds are produced results in optimal perception whereby listeners learn to categorize the [phonetic] speech signal into the [phonological] vowels and consonants that match the ones intended by the speaker" (Escudero, 2005: 52). The way in which the cue constraints are ranked (which types of cue constraints are looked at, in ascending order, when phonologically mapping a speech signal) are based on "distributions of the acoustic values with which sound categories are produced" (Escudero, 2005: 52). Not only does production have its basis in perception, Escudero also points out that perception must inevitably have its basis in production; she attaches considerable importance to what is intended by the producer of a speech signal: the Optimal Perception Hypothesis bears on "how likely it is that [phonetic] acoustic values were intended as a given [phonological] sound category" (2005:52).

In section 2.3.2, the key points of Escudero's L2LP (2005) will be assessed with the above concepts in mind.

2.3.2 Key points of the L2LP

Escudero asserts that L2 perception "is described using linguistic perception grammars and phonological categories" (2005: 86), which implies that there is an interaction as well as a

distinction between the phonetic and the phonological levels of analysis. In other words, a perception grammar helps parsing the (phonetic) auditory speech signal into segments which are then assigned to phonological categories. Escudero adds a nuance here that is essential to L2 perception: "a separation of perceptual mappings from sound representations [of different languages] leads to an adequate comparison of the perception systems involved" (2005: 86). Different languages have a different distribution regarding the boundaries of their cue constraints, even if these different cue constraints result in comparable phonological categories. As an illustration, Escudero presents a comparison of Canadian French (phonetic) [ɛ-æ] production and subsequent (phonological) /ɛ-æ/ perception with the Canadian English production and perception of the same vowel contrast (2005: 86-87). She observes here that "although the phonological categories in the two languages are the same, or are described with the same abstract symbols, the optimal perception hypothesis predicts that their perceptual mappings [i.e. the exact boundaries of the cue constraints] will be different" (Escudero, 2005: 87).

Escudero states that the first step in any explanation of L2 sound perception should therefore be "to conduct a thorough analysis of the optimal perception in each of the languages involved" (2005: 87). As stated before, the optimal perception hypothesis depends heavily on production environment. Accordingly, the optimal perception of any language is logically achieved by native listeners of that particular language - and not, instance, by L2 learners, as the required production environment is generally lacking in these listeners. It is for this reason that Escudero deems this optimal perception analysis of both the L1 and the target L2s to be absolutely necessary. Apart from the "location of the category boundaries [and] the shape of category boundaries", the "relative use of auditory dimensions" (i.e. the relative importance of different auditory cues) is important as well (Escudero, 2005: 90).

This description of the optimal perception of the L1 and the target L2 is only one of the five theoretical ingredients that constitute Escudero's L2LP (2005), and that "give an

explicit prediction, linguistic explanation, and phonetic/phonological description of L2 sound perception at the three logical states [viz. the initial, developmental and end state] of the acquisition process" (Escudero, 2005: 95). Figure 2.1 visualises the theoretical ingredients.

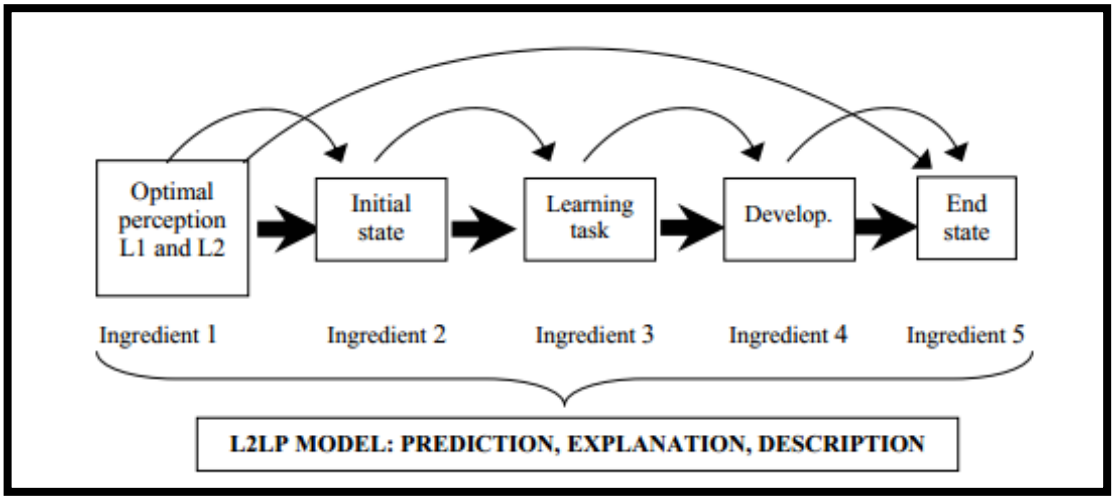


Figure 2.1. The L2LP theoretical ingredients (taken from Escudero, 2005: 95).

Following the optimal perception hypothesis (ingredient 1), the initial state (ingredient 2) ties in with the optimal L1 perception (initially fully copied), while the end state (ingredient 5) corresponds to the optimal perception of the target L2 (i.e. how it is perceived by an L2 native speaker). The learning tasks (ingredient 3) are the tasks involved in successfully taking the initial steps in the transition from initial to end state. These learning tasks depend on how the perception grammars of both languages differ from each other, and how these differences can be overcome. Fine-tuning of L2 perception takes place in the developmental stage (ingredient 4).

In what follows, the different L2LP (Escudero, 2005) ingredients will briefly be discussed. Table 2.1 below provides an overview with the prediction, explanation and description of each of Escudero's L2LP (2005) ingredients. The first ingredient has already been discussed extensively in the above paragraphs. In the second ingredient, it is predicted that "cross-language perception, i.e. the perception of a language by a listener with no prior

knowledge of that language, constitutes the L2 initial state" (Escudero, 2005: 98-99). The perception in this initial state could be compared to the non-native perception of naïve listeners in Best's PAM (1995). In this phase, 'full copying' of the L1 perceptual mapping takes place, and all of the L2 sounds are mapped onto L1 phonological categories, as if they were L1 sounds. The third ingredient, which constitutes the learning tasks, describes the different ways in which listeners try to "[r]each the optimal target L2 perception" by "[b]ridging mismatches between L1 and target optimal perception" (Escudero, 2005: 122). As mentioned before, Escudero's L2LP (2005) distinguishes clearly between perceptual mapping and phonological representation. Therefore, Escudero argues that "there are two types of learning tasks in L2 sound perception, namely a *perceptual task* and a *representational task*" (2005: 107). The perceptual task controls the way in which the optimal perception grammar of the L1 is adjusted to accurately process the L2 speech signal, while the representational task is concerned with the potential creation or reduction of new phonological categories.

The development of L2 perception by L2 learners is predicted to mirror that of L1 perception by children and infants. The L2 learner's full access to the GLA or Gradual Learning Algorithm (cf. Escudero, 2005: 68) plays an important role. This GLA will not be discussed here; the principal consideration to bear in mind here is that an essentially identical mechanism is responsible for L1 and L2 acquisition, and that development results in the formation of new categories and the shift of category boundaries (Escudero, 2005: 122).

In the end state, finally, "[o]ptimal L1 perception and optimal L2 perception" are predicted. As the explanation and description of this state involve cognitive mechanisms which will not be looked into any deeper, it suffices to note that in this end state of L2 perception acquisition, optimal L2 perception has been acquired, while optimal L1 perception is still maintained.

Table 2.1. A summary of the L2LP's theoretical ingredients (taken from Escudero, 2005: 122).

L2LP	Prediction	Explanation	Description
Optimal L1 & target L2	Human beings are optimal listeners	Optimal listeners handle the environment maximally well	L1 and L2 optimal category boundaries: Location & shape
Initial state	= Cross-language perception	Full Copying	L1 boundary location and shape
Learning task	= Reach the optimal target L2 perception	L2	Bridging mismatches between L1 and target optimal perception
Development	= L1-like	Full GLA Access	Category formation and boundary shifts
End state	Optimal L1 perception and optimal L2 perception	Input overrules plasticity Separate grammars	Language activation modes, through language setting variables

Methodologically, Escudero's L2LP model (2005) is considered useful with regard to this particular study, as its focus on production environment as essential to optimal L1 and L2 perception leaves sufficient room for dialectal differences to play a role in L2 perception (and the subsequent L2 production). Escudero indeed asserts that dialectal differences in the target L2 (e.g. in Southern British and Scottish English with regard to the perception by Spanish listeners, 2005: 20) can influence L2 vowel perception. In Escudero & Williams (2012) (discussed in section 3), it is proven that L1 (Spanish) dialect differences can have an influence on L2 (Dutch) vowel perception as well.

The following subsection discusses the different ways in which L2 sound categories relate to the L1 sound system.

2.3.3 Comparing L1 to L2 optimal perceptions: three scenarios

Escudero presents three types of scenarios that occur in the perception of an L2 by L2 learners, which differ depending "on how [their L1 perception grammar] compares to the optimal perception of different target L2s" (2005: 123). What follows is a brief overview, as the description of these scenarios and the way they tie in with Escudero's L2LP (2005) are methodologically useful for the present study. Table 2.2 presents the initial states and learning tasks for the three L2LP scenarios.

Table 2.2. Predicted initial states and learning tasks for the three L2LP scenarios (taken from Escudero, 2005: 125).

	NEW	SUBSET	SIMILAR
Initial state	Too few categories	Too many categories	Non-optimal mappings
Perceptual task	Two tasks: Creation and integration	One task: Category boundary shift	One task: Category boundary shift
Representational task	Two tasks: Create features and turn them into segments	Two tasks: Reduce lexical and perceived categories	none
Relative difficulty	Most difficult	Medium difficulty	Less difficult

In the NEW scenario, "the L2 environment produces phonological differences that do not exist in the L1" (Escudero, 2005: 123). Creation and integration of new mappings account for a double perceptual (phonetic) task (Escudero, 2005: 125). The representational tasks help identify NEW sounds as distinct phones (Escudero, 2005: 125). That this scenario is relatively most difficult is "mainly because the NEW scenario not only involves the *creation* of new categories and perceptual mappings but also the integration of the new categorized dimensions with already categorized dimensions" (Escudero, 2005: 125).

In the SUBSET scenario, "the L2 categories constitute a SUBSET of [the] L1 categories" (Escudero, 2005: 123). Perceptually, this implies an adjustment of the different category boundaries (no creation and integration is needed here, as the SUBSET L2 categories overlap acoustically with L1 categories). On a representational level, the perceived phonological categories have to be reduced to achieve accurate L2 perception. SUBSET scenarios are regarded as relatively less difficult than NEW scenarios, as only one perceptual task is involved (Escudero, 2005: 125).

The third scenario involves L2 sounds that are SIMILAR to L1 sounds, which occurs when "the L1 and L2 categories are equivalent" (Escudero, 2005: 123). Non-optimal mappings of L2 sounds on L1 perceptual maps are the result in the initial state. In contrast with NEW and SUBSET scenarios, where there are too few and too many categories, respectively, the SIMILAR scenario has equivalent L1 and L2 categories: no representational tasks are needed, and the only perceptual task is a category boundary shift to optimally perceive the relevant L2 sounds (Escudero, 2005: 125). Therefore, this is seen as the least difficult of the three scenarios Escudero proposes (2005: 125).

To bring the present section to a close, section 2.4 will provide a brief comparison of Escudero's L2LP model (2005) with Flege's SLM (1995) and Best (/& Tyler)'s PAM(/-L2) (1995/2007), the two previously discussed models.

2.4 SLM, PAM-L2 and L2LP: a comparison

In a comparison of the L2LP (2005) with previous theoretical models, Escudero also includes Flege's SLM (1995) and Best's PAM (2007). Escudero's dissertation dates from 2005, so she can not yet have included an assessment of the 2007 PAM-L2; however, as asserted by Best & Tyler, "L2 perceptual learning is determined by non-native speech perception principles" (2007: 2), so that a critical assessment of how Escudero evaluates Best's PAM (1995) is justified and relevant to the present study.

Escudero uses the framework of the L2LP (2005), with its attention not only to the initial state, development and end state of L2 perception, but also to speech perception itself, L1 perception and the possible L2 sound perception scenarios. This review will consider those elements of Flege's SLM (1995) and Best's PAM (1995) that contrast, but also correspond with Escudero's L2LP (2005). Table 2.3 below is used to schematically present the three theoretical models and how they differ and correspond on some key points that are assessed in all three models.

Table 2.3. Comparison of SLM, PAM(-L2) and L2LP. The table is based on Flege (1995), Best (1995), Best & Tyler (2007) and chapter 4 of Escudero (2005, 127-153).

	SLM (Flege, 1995)	PAM(-L2) (Best, 1995/ Best & Tyler, 2007)	L2LP (Escudero, 2005)
General speech perception	• No distinction between perceptual mapping and sound representations • Seemingly conflated: "stored categories are seen to perform the mapping of the signal" (Escudero, 2005: 131) • Phonetic categories as mental representations	• Distinction between perceptual mapping and sound representations, but "[a] common gestural domain for both phonetic details and phonological structure" (Best, 1995: 183) • Phonological categories comprised of sets of gestural features	• Distinction between (phonetic) perceptual mapping and (phonological) sound representations • Perception grammars help parsing phonetic auditory speech signals into segments which are then assigned to phonological categories
L1 acquisition	• Emphasis on L1 production environment for language-specific sound perception (no innate phonetic/phonological categories) • Escudero's critique (2005: 132): L1 learning mechanisms are left undefined (except perhaps the concept of equivalence classification)	• Emphasis on L1 production environment for language-specific sound perception (no innate phonetic/phonological categories) • Escudero's critique (2005: 132): "the learning mechanism that allows the infant to 'pick up' high level articulatory gestures is left undefined"	• Similar emphasis as SLM and PAM • "[T]he emergence of language-specific sound perception is the result of exposure to a particular language environment" (Escudero, 2005: 131) • Psycholinguistic/cognitive explanations for L1 learning (not discussed in the present study)

Table 2.3. Comparison of SLM, PAM(-L2) and L2LP (continued).

	SLM (Flege, 1995)	PAM(-L2) (Best, 1995/ Best & Tyler, 2007)	L2LP (Escudero, 2005)
L2 perception: initial state	• L1 phonetic categories essential to L2 perception: "Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 or L2 phones identified as a realization of each category" (Flege, 1995: 239) • Methodological critique by Escudero: fact that L2 perception should always be compared to perception of "absolute L2 beginners" not attended to (2005: 138)	• Initial L2 perception state = non-native perception of naive listeners • L2 perception of naive listeners depends on L1: "the relative ease or difficulty of a given contrast varies according to the listener's native language" (Best & Tyler 2007: 7) • Methodological critique by Escudero: fact that L2 perception should always be compared to perception of "absolute L2 beginners" not attended to (2005: 138)	• L2 initial state ~ nonnative cross-language perception (Escudero, 2005: 98) • "Full Copying of L1 perceptual mappings" & "phonemic equation of L2 and L1 sounds" (Escudero, 2005: 98) • "L1 optimal boundaries and categories constitute the initial L2 perception" (Escudero, 2005: 98) • L2 perception should always be compared to perception of "absolute L2 beginners" (Escudero, 2005: 138)
L2 perception: development	• L2 listener has access to same mechanism acquired for L1 perception	• L2 listener has access to same mechanism acquired for L1 perception	• Cognitive substantiation applied to assert that L2 development is very much L1-like (Escudero, 2005: 142).

Table 2.3. Comparison of SLM, PAM(-L2) and L2LP (continued).

	SLM (Flege, 1995)	PAM(-L2) (Best, 1995/ Best & Tyler, 2007)	L2LP (Escudero, 2005)
L2 perception: end state	• Age and production environment play essential role, but native-like L2 perception (and production) is possible • Common phonological space of L1 and L2: potential influence of L2 on L1	• End state not addressed • Influence of L2 on L1 is possible	• Optimal (native-like) perception of L2 is possible • Influence of L2 on L1 unlikely: separate perception grammars

L2 perception: scenarios	• Three-way distinction: *new*, *similar* and *identical* L2 sounds
	• *New* L2 sounds easy to learn (no interference from L1)
	• *Identical* L2 sounds easy too (positive L1 transfer)
	• *Similar* L2 sounds difficult (equivalence classification)
		• Works with new and similar sounds
		• Scenarios for contrasts relevant to this study:
		  ○ Only one L2 phonological category is perceived as equivalent to given L1 phonological category
		  ○ Category goodness assimilation contrast
		  ○ Single-category L2 contrast assimilation (L2 phonological categories perceived as equally good/poor instances of L1 phonological category
			• Two innovations:
			  ○ Scenarios are considered at each L2 perception stage
			  ○ New SUBSET scenario
			• Three scenarios
			  ○ NEW: too few categories
			  ○ SUBSET: too many categories
			  ○ SIMILAR: non-optimal mappings
			• Difficulty depends on amount of learning tasks: NEW > SUBSET > SIMILAR

3 The role of L1 regional variation in L2 perception and production[6]

3.1 Introductory notes

As discussed in the previous section, the linguistic production environment is vital for the perception and subsequent production of L1 and L2 languages. This finding implies that dialect can play an essential part not only in the perception and production of an L1, but also of an L2. As the saying goes, a dialect is a language lacking an army and a navy; certain regional features could persist in the daily language use. It is not unlikely that these features have an influence on the different perception and production of the same L2 by L1 speakers with different regional backgrounds.

It is possible that this regional influence also applies to the Flemish situation. Standard Dutch is the standard variety in Flanders. Although Standard Dutch is the prescribed norm in Flemish education, the reality is different. Van de Velde argues that the majority of Flemish people – adolescents included – has problems with speaking and writing Standard Dutch fluently and correctly (2002: 136, my translation)[7]. Moreover, he argues that Standard Dutch is used almost exclusively in formal situations, and that dialects are, compared to our neighbouring countries, still in a strong position and display an enormous linguistic diversity (Van de Velde, 2002: 137, my translation)[8]. These observations imply a Flemish production environment that is indeed shaped by regional influences. An

[6] Some of the studies discussed in the present section were also discussed in Debaene (2012); it should be pointed out that some parts are adapted from this BA paper. However, the primary sources are the same, and the parts in question are mostly descriptive in nature. Therefore, Debaene (2012) will not always be referred to as such if this is not relevant or essential. This also applies for the description of the vowel systems in section 4, and the description of the production experiment in section 6.2.

[7] "Wel kunnen we vaststellen dat de overgrote meerderheid van de Vlamingen - met inbegrip van de jongeren - problemen heeft om vlot en correct Standaardnederlands te spreken en te schrijven."

[8] "[...] het Standaardnederlands wordt bijna uitsluitend in formele situaties gebruikt; de dialecten staan - in vergelijking met de ons omringende landen - nog altijd sterk en vertonen een enorme linguïstische diversiteit [...]."

assessment of the influence of these regional characteristics on L2 sound perception and production, or even a point of view in which the regional variety is seen as the L1, is therefore predicted to yield fruitful results. This is especially the case in this particular study with its focus on vowels. Vowels are methodologically interesting, as "these segments are produced with higher intensity, longer duration and more acoustic dimensions than consonants" (Escudero, 2005: 19). Moreover, "[t]hey are also [generally] fewer in number, which makes them much more variable than consonants among languages *and even among dialects*" (Escudero, 2005: 19, my italics). Due to this extensive dialectal variability, L2 sounds in general and vowels in particular could be differently perceived and produced by L1 speakers from different regional backgrounds. What is, in Flege's (1997) terms, a *new*, *similar* or *identical* L2 sound to one group of dialect L1 speakers, does not necessarily have to be so for L1 speakers with a different L1 regional variety.

The insight that regional varieties of a native language can exert influence on L2 perception and production has only recently led to numerous studies which implement L1 (or L2) regional variety as a factor to be researched. In what follows, a concise review will be given of several studies with precisely this starting point. In section 3.2, studies that focus on perception will be looked at. In section 3.3, production studies are under consideration. Special attention will be given to how the theoretical frameworks discussed in section 2 are used to account for the results.

3.2 Influence on L2 perception

In her doctoral dissertation, Marinescu examined "the role of the native dialect in non-native perception and production in the specific case of Cuban and Peninsular Spanish as native varieties and of English vowels /æ, ʌ, ɑ/ as the target" (2012: ii). Methodologically, this study is interesting in that it uses the same approach as the present study. An acoustic assessment of the cross-dialectal L1 differences between the two varieties of Spanish precedes the investigation of L2 English perception and production. Moreover, not only spectral differences were discovered between the two dialects, but also durational ones, another element that is predicted to differentiate Antwerp (Brabantine) Dutch from Ghent (East-Flemish) Dutch. Although Marinescu found "little evidence for L1 dialect effect in L2 [perception]" (2012: 81), "[s]harper differences between the two groups of L2 learners with distinct dialectal backgrounds were highlighted [for L2 production], suggesting different interlanguage strategies of producing the English /æ, ʌ, ɑ/ vowels" (2012: 109).

In another study, Chládková & Podlipský (2011) acknowledge the importance of L2 variation in the process of second language acquisition (SLA). One of the studies they refer to is Escudero & Boersma (2004) (also mentioned in the previous paragraph). In this study, it was found that Spanish learners of English display a different L2 vowel perception based on dialectal differences in their target L2 language (i.c. Southern British and Scottish English). However, Chládková & Podlipský point out that in L2 vowel perception studies, L1 speakers have generally been pooled together; they argue that "[n]ot taking into account learners' specific L1 dialect background could have obscured possible systematic differences between learners" (2011: 187). The authors specifically investigated the influence of dialectal differences in Bohemian and Moravian Czech on the perception of Dutch vowels. The authors focus on the non-low non-back vowel region, in which "Czech has fewer vowel phonemes than Dutch [...] (four as opposed to seven)" (2011: 188). This leads them to predict, in line with Best's PAM (1995) and Escudero's L2LP (2005), "a number of single-

category assimilations [...] for vowels in this region". What is particularly interesting in this study is that, as in Marinescu (2012), the durational and spectral vowel qualities differ with regard to the two different dialects: "Moravians, but not Bohemians, favor the durational difference over the spectral difference when perceiving the Czech /i:-ɪ/ contrast" (2011: 188). Dutch (at least the Standard variant) has a predominantly spectral difference in its /i-ɪ/ contrast. The authors therefore predict different perceptual assimilations by the two groups of Czech speakers:

> *"Moravian listeners are expected to assimilate both of these phonetically short non-native vowels to a single native category, namely /ɪ/, whereas Bohemian listeners are more likely to perceive this Dutch vowel contrast in terms of two native categories, namely, /i:/ and /ɪ/"* (2011: 188).

The results of the study supported Escudero's L2LP model (2005), since they revealed that "Bohemian Czech listeners differed from Moravian Czech listeners in how the Dutch /i-ɪ/ and /y-ʏ/ contrasts were perceived in terms of L1 categories" (2011: 191). In their conclusion, Chládková & Podlipský point out that "[a]lthough it may seem to be a trivial requirement, it is by no means standard practice in cross-language and L2 speech perception studies to consider not only L1, but also L1 dialect as a factor affecting listener's performance" (2011: 191-92).

In their 2012 study, Escudero & Williams examined how Peruvian and Iberian Spanish learners of Dutch performed in categorical discrimination and identification tasks of Dutch vowels. The prediction was "that the acoustic differences between the vowel productions of the two dialects, which compare differently to Dutch vowels, would manifest in differential L2 perception for listeners of these two dialects" (Escudero & Williams, 2012: 406). Both of the Spanish varieties under consideration have only one /i/ vowel in the vowel space in which Dutch has both /i/ and /ɪ/. The realisations of /i/ in Peruvian and

Iberian Spanish do, however, display acoustic differences that could play a role: Peruvian Spanish /i/ lies acoustically in the middle between Dutch /i/ and /ɪ/; the expected result is that these two Dutch vowels are perceptually assimilated equally to Peruvian Spanish /i/ - and subsequently discriminated poorly (Escudero & Williams, 2012: 408). Following Best's PAM (1995), Best & Tyler's PAM-L2 (2007) and Escudero's L2LP (2005), the authors remind us that "non-native contrasts that are mapped on to a single native category lead to the most difficulty in discrimination and learning" (2012: 406). Iberian Spanish /i/, on the other hand, is acoustically closer to Dutch /ɪ/ than to Dutch /i/. This gives rise to the slightly different PAM scenario with a *Category-Goodness Difference*: the two sounds of the contrast are "assimilated to the same native category, but they differ in discrepancy from native "ideal" [i.c. Spanish /i/]" (Best, 1995:195).

Another contrast under consideration in Escudero & Williams (2012) was Dutch /a-ɑ/, which was also predicted to be easier to Iberian Spanish listeners. This is because this particular contrast is acoustically closer to the Iberian Spanish /a/ and /o/ vowels (to which the vowels of the Dutch contrast are likely to be perceptually assimilated) than to Peruvian Spanish /a/ and /o/. The results supported the predictions of Escudero & Williams: "[Iberian Spanish] learners were more accurate than PS learners in discriminating the two most difficult contrasts, i.e., /a-ɑ/ and /i-ɪ/", suggesting that "acoustic differences in the production of [Iberian Spanish] and [Peruvian Spanish] vowels lead to a difference in L2 vowel perception" (2012: 411). This study thus once more confirms Escudero's L2LP (2005), with its focus on different acoustic properties and its reliance "on the comparison of [these] acoustic properties between dialects and languages" (Escudero & Williams, 2012: 411).

To conclude this subsection on L2 perception, a study by Escudero, Simon & Mitterer (2012) is considered. In this study, the influence of regional variation in L1 Dutch (North Holland Dutch as opposed to East- and West-Flemish Dutch) was examined. Specifically, it

was investigated how listeners from these different regional backgrounds perceived the English /ɛ - æ/ contrast. Although this contrast is absent in both the Holland and the Flemish varieties, those two varieties "differ in the phonetic realisation of the Dutch front vowel /ɛ/" (Escudero, Simon & Mitterer, 2012: 281). Methodologically, Escudero's L2LP (2005) is followed, with its proposal that a detailed analysis of the relevant L1 sounds as well as the relevant L2 sounds can predict L2 perception quite accurately. Specifically applied to this particular study, the authors examined, among other thing, "to what extent acoustic similarity between Dutch /ɛ/ in different varieties of Dutch and the two English front vowels can predict patterns of perceptual assimilation" (Escudero, Simon & Mitterer, 2012: 281). One of the conclusions was that the two groups of Dutch listeners "differed in their non-native perception of the English vowels /ɛ/ and /æ/" and that "dialectal variation in the acoustic properties of Dutch /ɛ/ seems to account for differences in non-native vowel perception" (2012: 286). Moreover, the data in this study "show that the exact acoustic vowel properties in the variety of the native language can predict how vowels in a second language will be perceived." This conclusion, with its focus on an acoustic/phonetic rather than an abstract phonological level, is in line with both Escudero's L2LP (2005) and PAM-L2, but it nevertheless leaves the question "as to whether the critical features for similarity across comparisons are the acoustic properties themselves (as assumed by L2LP), or the articulatory gestures which give rise to them (as assumed by PAM-L2)" (2012: 287).

3.3 Influence on L2 production

In the previous subsection, it was already mentioned that Marinescu (2012) not only investigated the influence of dialectal L1 differences on L2 perception, but also on L2 production. What is relevant for the present study, is that Marinescu found that "the L2 production experiment [...] refined the findings of the L2 perceptual experiment", and that "[s]harper differences between the two groups of L2 learners with distinct dialectal backgrounds were highlighted, suggesting different interlanguage strategies of producing the English /æ, ʌ, ɑ/ vowels" (2012: 109).

In a study by Robert Lew it was investigated "if, and how, regional differences in Polish with respect to regressive sandhi voicing assimilation may influence the corresponding aspects of learners' English pronunciation" (2002: 1). Lew points out that "regressive sandhi voicing before consonants is very rare in native English speech but is a noticeable feature of a foreign accent" (2002: 2). This study is one of the few studies that considers consonants, which, as already mentioned in the introduction to the present section, are less likely to differ across dialects (and languages) than vowels. It would strengthen the validity of L1 regional influence on L2 if the regional characteristics of L1 consonants could also have an influence on the L2. Lew provides support that this is indeed the case. In Polish, obstruent clusters - even across lexical and (some) morphological boundaries - correspond in voicing: "the voicing feature of the onset obstruent or obstruents of the following word (morpheme) will determine the voicing of the word-final obstruent or obstruents in the preceding word (morpheme)" (2002: 1). In the case of the second element in such a cluster being a non-obstruent, however, regional variation is observed in the (de)voicing of any preceding word-final obstruent(s): in some areas (namely Mazowsze and Pomerania) 'devoicing' varieties of Polish are spoken; in other areas (Silesia, Wielkopolska, and Małopolska), 'voicing' varieties of Polish are found (cf. examples 1 and 2, respectively) (2002: 1).

(1) <kot/kod albo> [k ɔ t a l b ɔ]

(2) <kot/kod albo> [k ɔ d a l b ɔ]

This regional variation is interesting since it is concerned with a contrast that - although it manifests itself in different regional allophones on a phonological level - can only be seen if a higher, morphological and cross-lexical level of analysis is considered. This could prove to be a fruitful line of future investigation. Lew found that this Polish regional variation indeed had an effect on the production of English: he discovered that in pre-sonorant (i.e. non-obstruent) contexts "[participants from voicing areas] voice over three times as frequently as [participants from devoicing areas]" (2002: 10).

O'Brien & Smith (2010) examined the production of German /u:/ and /y:/ by speakers from three North American English dialect regions (the Inland Northern and North Central dialects of the U.S., along with Western Canadian English) and rely on Flege's SLM (1995) to account for the results. The authors considered three research questions. Methodologically following Escudero's L2LP (2005) guideline of a thorough acoustic analysis of the relevant sounds in the L1 dialects, they first examine how speakers from the different North American dialect regions produce /u:/, since a great deal of regional variation can be found in the realisations of this vowel. A second research question addresses whether "subjects [will] also differ in their production of the German vowels /u:/ and /y:/ based on their dialect of North American English" (2010: 306). As German /y:/ is a nonnative vowel to English speakers, it could, in the terms of Flege (1995), be seen as *new* to English speakers, and thus relatively easy to perceive and produce. However, the authors ask themselves if "subjects from Western Canada and the Inland North whose English /u/ has been claimed to be an allophonic variant of [y] [will] even be able to establish a new category for this vowel" (2010: 307). This is a very interesting case: not only are there dialectal differences between the L1 productions of /u:/ which could influence the L2 German /u:/, but there is the extra factor of the non-native /y:/, which is roughly situated in the same vowel region

as /u:/, spectrally. The authors found that speakers from the three dialect regions did not assimilate German /u:/ to their English /u/. However, no major dialectal influence was discovered, since "subjects in all three dialect regions produced their German /u:/ with significantly higher F2 values (suggestive of more fronting) than their English /u/" (O'Brien & Smith, 2010: 320). Speakers from all three dialect regions were able to create a new category for the German /y:/; "however, they differed in how they ultimately established the German /u:/-/y:/ contrast", in that learners from the two U.S. regions used the F2 formant value to contrast these vowels, while Western Canadian learners used F3 (O'Brien & Smith, 2010: 297; F3 is a formant indicative of lip rounding, cf. O'Brien & Smith, 2010: 299). The authors concluded that "L2 vowel formant values differ by dialect region even when the learners' L1 dialects differ only subtly" (2010: 298). This conclusion once more supports Escudero's L2LP (2005), which attaches a lot of importance to acoustic properties in its L2 research.

In conclusion, this section shows the scientific relevance and value of considering the factor of L1 regional difference in SLA research, as all of the (very recently conducted) studies mentioned here proved to some extent that there is indeed an influence of this factor on L2 perception and production. This overview looks especially promising when considering dialects of Flemish Dutch, as regional variation is still relatively strong in Flanders. Another fact to be remembered from the present section is the methodological value and necessity of carefully examining the phonological and acoustic features of the L1 dialects and the L2. The assessment of the vowel systems relevant to the present study, namely Standard Dutch, Ghent East-Flemish and Antwerp Brabantine Dutch, along with RP British English, will be dealt with in the next section.

4 Vowel systems in contrast

Preliminary notes

Following the methodology prescribed in Escudero's L2LP (2005), the present section provides an overview of the vowel systems of the L1 and L2 varieties relevant to the present study, namely L1 Standard Southern Dutch, Ghent East-Flemish Dutch and Antwerp Brabantine Dutch, along with L2 RP English. This section will not consider the entire vowel system of each variety under consideration, but will only consider the vowels that are relevant to the present study, i.e. those in the L1 Dutch /i-ɪ/ and /y-ʏ/contrasts, and the RP English vowels /iː-ɪ/, /ʌ/ and /ʊ/. These last two vowels are selected as they are arguably spectrally closest to the Dutch /y/ and /ʏ/, for which there are no (in Flege's 1995 SLM terms) *similar* English counterparts. As in Debaene (2012), the British variety was chosen over the American English variety, since all participants indicated to have been taught British RP English in secondary school. (cf. Appendix A). Moreover, Collins & Mees (2003: chapter 28) illustrate that none of the differences between the two varieties apply to the vowels that are relevant to the present study (except maybe a generally closer realisation of /ʌ/ in American English).

4.1 L1 vowel systems

4.1.1 Southern Standard Dutch

It is not until the studies conducted by Verhoeven & Van Bael (2002a&b) and Adank, van Hout & Smits (2004a) that Southern Standard Dutch (i.e. the variety spoken in Flanders) vowels were investigated systematically. Indeed, Verhoeven & Van Bael point out that "[t]he interest for vowels in Flanders has been restricted to a few impressionistic descriptions which often provide pronunciation guidelines for dialect speakers who want to improve their pronunciation of Standard Dutch" (2002a: 150). As Verhoeven & Van Bael (2002a&b) use a regional approach, the results of their study will be assessed in sections 4.1.2 and 4.1.3. In this section, the focus is on studies that use acoustic measurements as their data (wherever possible). However, because a detailed and statistical assessment of the studies is not possible due to different vowel normalisation methods or consonantal contexts, attention will primarily be paid to how the vowels under consideration relate to each other on a durational and a spectral basis.

The study of Adank *et al.* compared the vowels of Northern and Southern Standard Dutch produced by teachers of Dutch at secondary education institutes, as these speakers "are expected to speak Standard Dutch on a daily basis" and can be "regarded as having a normative role" (2004a: 1730). The four Dutch vowels under consideration in the present study are close front unrounded /i/, the contrasting near-close near-front unrounded /ɪ/, close front rounded /y/ and near-close near-front rounded /ʏ/. Durationally, these vowels belong to what the authors refer to as "the group of [relatively seen] shorter vowels" (Adank *et al.*, 2004a: 1737), which is consistent with the measurements of Nooteboom, who concludes that /y/ and /i/ are short vowels (1971: 401). Other studies which are mentioned by Adank *et al.* (2004a) (namely Koopmans-van Beinum, 1980 and Rietveld *et al.*, 2004) classify /ɪ/ and /ʏ/ in the group of short vowels, and /i/ and /y/ in the group of half-long vowels (yet never in the group of long vowels).

On a spectral level, "the nine monophthongal vowels of Dutch, /ɑ a ɛ ɪ i ɔ u ʏ y/, could be separated fairly well based on their steady-state characteristics for their first two formant frequencies alone" (Adank *et al.*, 2004a: 1737). The only spectral and durational differences between Northern and Southern Standard Dutch that were revealed did not apply to monophthongs, which were found to be fairly consistent in the two varieties. Therefore, Leussen, Williams & Escudero (2011) is also considered, in which contextual effects on Dutch monophthongs were examined, along with a comparison to previous studies. When considering Table 4.1, which lists the average acoustic values for the corpus used in the study of Leussen *et al.* (2011), it appears that vowel duration does not seem to play a differentiating role in the /i-ɪ/ and the /y-ʏ/ contrast.

Table 4.1. Acoustic values for /i ɪ y ʏ/ in the corpus of Leussen *et al.* (2011), logarithmically averaged over all male (M)/ female (F) tokens (adapted from Leussen *et al.*, 2011).

		F1 (Hz)	F2 (Hz)	F3 (Hz)	Duration (ms)
/i/	F	354.58	2644.04	3229.29	90.86
	M	316.85	2127.72	2782.6	90.46
/ɪ/	F	460.55	2275.89	2922.68	89.33
	M	405.18	1852.35	2566.99	87.55
/y/	F	372.71	2031.29	2731.46	95.91
	M	325.15	1697.23	2264.94	95.55
/ʏ/	F	462.79	1813.64	2673.17	91.88
	M	420.12	1498.84	2225.88	89.88

Figure 4.1 maps the F1 and F2 values for Standard Dutch male monophthongs from three different studies: Pols *et al.* (1973), Adank *et al.* (2004a), and the study of Leussen *et al.* (2011) itself[9]. Despite considerable differences between studies, it can be observed that the manner in itself in which the relevant contrasts spectrally relate to each other as such remains considerably constant across the three studies.

[9] Figure 4.1 uses a representation based on the Bark Difference Metric, a vowel normalisation method).

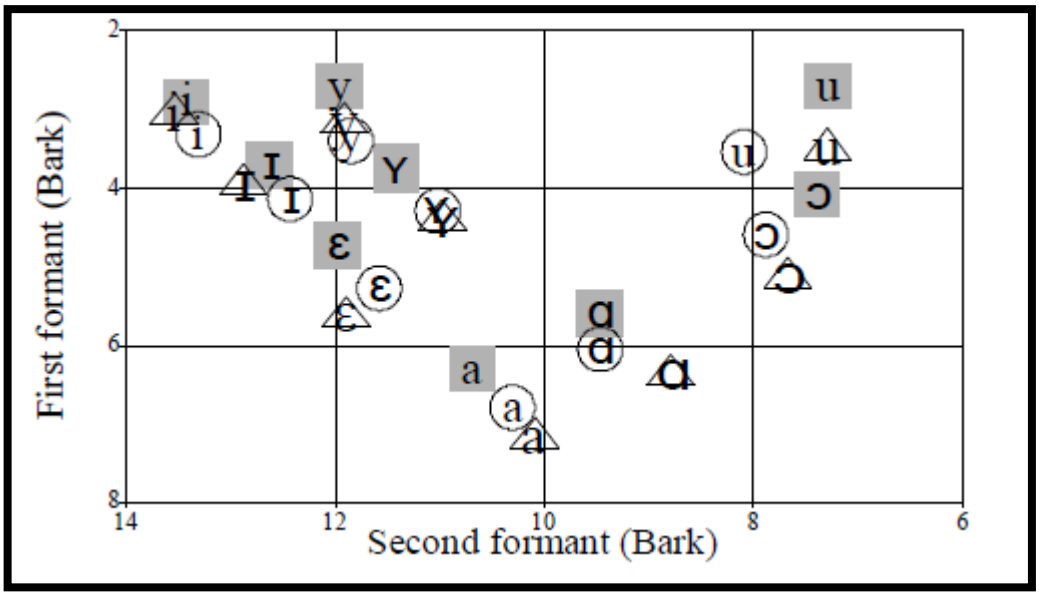


Figure 4.1. Average positions of male vowels on the F1/F2 plane for data from Pols *et al.* (1973) (triangles), Adank *et al.* (2004a) (rectangles) and Leussen *et al.* (2011) (circles) (taken from Leussen *et al.*, 2011: 3).

In conclusion, the previous results reveal that the Standard Dutch monophthongs are produced the same way in the Northern and the Southern variety. They furthermore suggest that the distinction between the /i-ɪ/ and the /y-ʏ/ contrast is made on the spectral rather than on the durational level in Standard Dutch.

4.1.2 Ghent (East-Flemish) Dutch

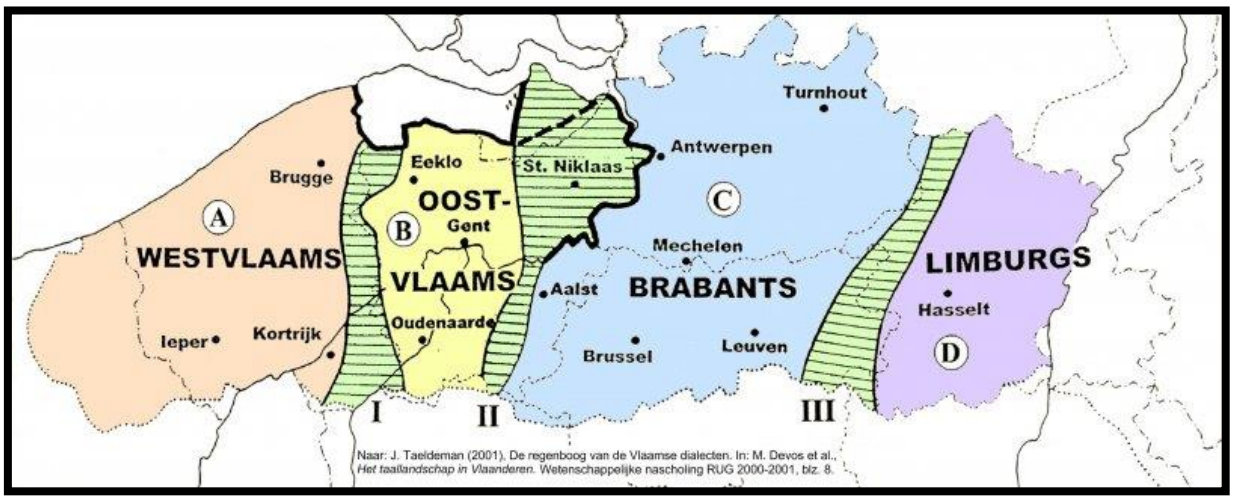


Figure 4.2. The dialect regions of Flanders: (from left to right) the West-Flemish, East-Flemish, Brabantine dialect, and Limburgish (taken from Triest, 2009a).

As can be seen on the map in Figure 4.2, the Ghent dialect belongs to the group of East-Flemish dialects. However, it is known for its dialect island character (Taeldeman,

1999). It shares with the surrounding East-Flemish dialects the omission of the durational distinction in vowels, so that, As Taeldeman observes, historically short and historically long vowels are pronounced – at least in the same phonological environment - with the same [durational] quantity (1985: 39, my translation)[10]. While in the other East-Flemish dialects this results in a shortening of the long vowels, in the Ghent dialect, the short vowels become longer (Taeldeman, 1985: 39). As in Standard Dutch, in the Ghent dialectal variety of Dutch /i-ɪ/ and /y-ʏ/ are thus differentiated on a spectral rather than on a durational level.

However, this does not mean that we can expect our participants to produce these vowels very long: first, Taeldeman points out that "the longer the monopthongs, the more "vulgar" the Ghent dialect (and vice versa)" (1985: 39) – as the participants of the present study are young students, they are not expected to speak with a broad accent and may be influenced by the standard language. In any case, this does not change the observation that there is a spectral rather than a durational distinction between the relevant vowel contrasts.

Although the East-Flemish vowels in Verhoeven & Van Bael (2002 a&b) are produced by speakers from Eeklo (approximately 20 km to the north-east of Ghent), the study is worth considering as it actually measured spectral and durational qualities of the vowels (Taeldeman, 1985 does not base his work on acoustic measurements). Their spectral data are consistent with what has previously been discussed in this section (Verhoeven & Van Bael, 2002a: 157). Moreover, Figure 4.3 reveals that the durational difference between /i/ and /ɪ/ is relatively small (they are also both classified as short vowels). However, what is striking is that East-Flemish /y/ has roughly double the duration of East-Flemish /ʏ/. Verhoeven & Van Bael, in the discussion of their results, furthermore point out that "in the

[10] "historisch korte en historisch lange vokalen worden er - althans in dezelfde fonologische omgeving - met dezelfde kwantiteit gerealiseerd."

East-Flemish variety the durational difference between long and short vowels is bigger [than in the Limburg variety]: short vowels are extra short, while long vowels are extra long" (2002a: 161). This is remarkable, "since the East-Flemish dialects are generally characterised by a neutralisation of long and short vowels" (Verhoeven & Van Bael, 2002a: 161), as already discussed above. This neutralisation can be found in the /i-ɪ/ contrast, yet not in, among others, the /y-ʏ/ contrast. The authors suggest that "the strong difference between short and long vowels in East-Flemish Standard Dutch may have to be accounted for in terms of hypercorrection" (2002a: 161).

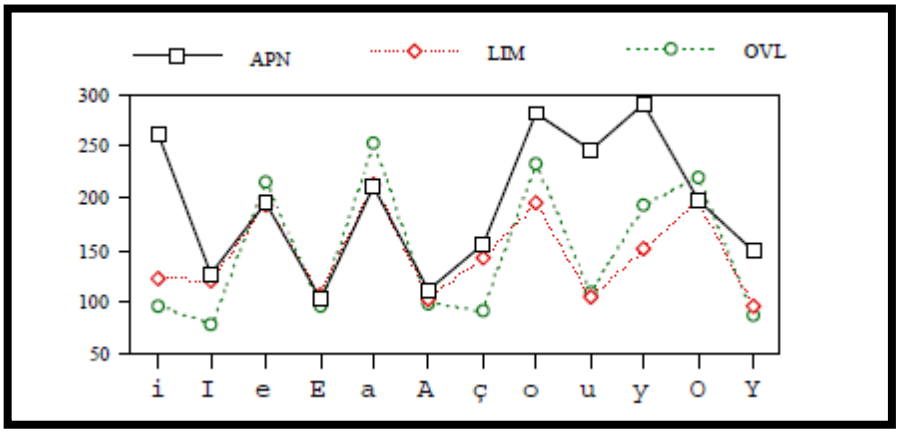


Figure 4.3. Average vowel duration (in ms) of the 12 Dutch vowels in Antwerp (APN), Limburg (LIM) and East-Flanders (OVL) (Adapted from Verhoeven & Van Bael, 2002a: 159).

Verhoeven & Van Bael additionally point out that "the acoustic characteristics of the Limburg and East-Flemish vowels are in good agreement, while the vowels of the Antwerp variety are considerably different" (2002a: 161). The fact that two quite separate regions as East-Flanders and Limburg (cf. Figure 4.2) display similar acoustic characteristics could suggest a top-down Standard Dutch influence. After all, the authors investigated regional varieties of Standard Dutch (and specifically designed their tasks for this purpose). The durational discrepancy between East-Flemish /y/ and /ʏ/ could therefore also be the result of Standard Dutch influence, in which these vowels have in some studies been characterised as 'half-long' and 'short', respectively (cf. section 4.1.1). All things considered, the vowel

production of Ghent participants is therefore hypothesised to be rather similar to that of Standard Dutch both in terms of durational and of spectral features.

4.1.3 Antwerp (Brabantine) Dutch

The Brabantine dialect area roughly coincides with the Dutch province of North Brabant and the Belgian provinces of Antwerp and Flemish Brabant. This discussion focuses on the Brabantine variety spoken in the city of Antwerp and its vicinity.

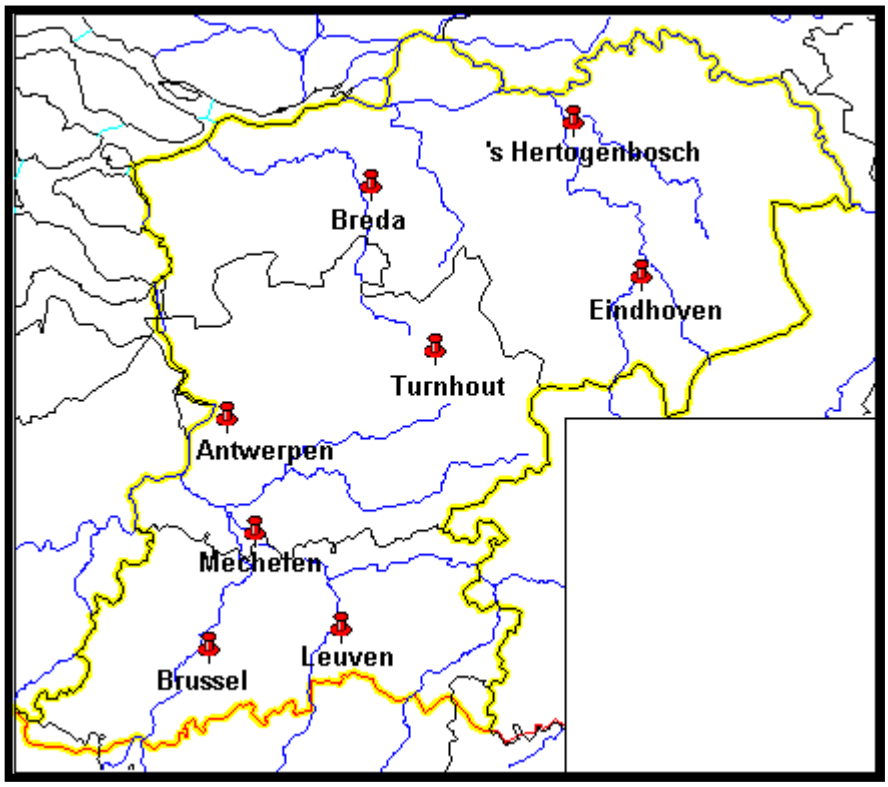


Figure 4.4. The Brabantine dialect area as it is treated in the *Dictionary of Brabantine Dialects* (taken from Keulen, van de Wijngaard & Van Keymeulen, 2010).

The section of the *Dictionary of Brabantine Dialects* that concerns sound geography (Belemans, Weijnen, Van Hout & Van Langendonck, 2000: 113) points out that the Standard Dutch short /ɪ/ vowel is problematic for many (especially Southern) Brabantine speakers. In this regional variety of Dutch, the /ɪ - i/ contrast is differentiated on another basis. The tendency is towards a longer realisation of the Standard Dutch short /i/ (which in Standard Dutch is only long preceding an /r/, cf. Nooteboom, 1971: 397), and a more closed,

somewhat fronted realisation of the short vowel /ɪ/, resulting in a durational rather than a spectral distinction in this vowel contrast. Nuyts goes as far as to transcribe the Antwerp realisation of Standard Dutch /ɪ/ with /i/, and the Antwerp realisation of Standard Dutch /i/ as a long /ɪː/ (1989: 24). While the transcription of /ɪ/ by /i/ and the length marks seem justified (cf. Figure 4.3), it is unclear what could justify the use of the symbol /ɪː/ to transcribe /i/, especially when considering Figure 4.5 below; Antwerp /i/ is produced almost identical to Limburg /i/, and lies spectrally far from the Limburg and East-Flemish productions of the vowel normally transcribed as /ɪ/. Thus, it would have seemed more correct for Nuyts (1989) to transcribe the Antwerp variety of Dutch /i/ as /iː/.

The Standard Dutch /y/ vowel is transcribed in Nuyts with a length mark (/yː/) and the Standard Dutch /ʏ/ vowel is transcribed as /y/ (1989: 24). The acoustic measurements by Verhoeven & Van Bael (2002a) support Nuyt's observations and transcription method, as they reveal a durational rather than a spectral distinction for this vowel pair (cf. Figures 4.3 and 4.5).

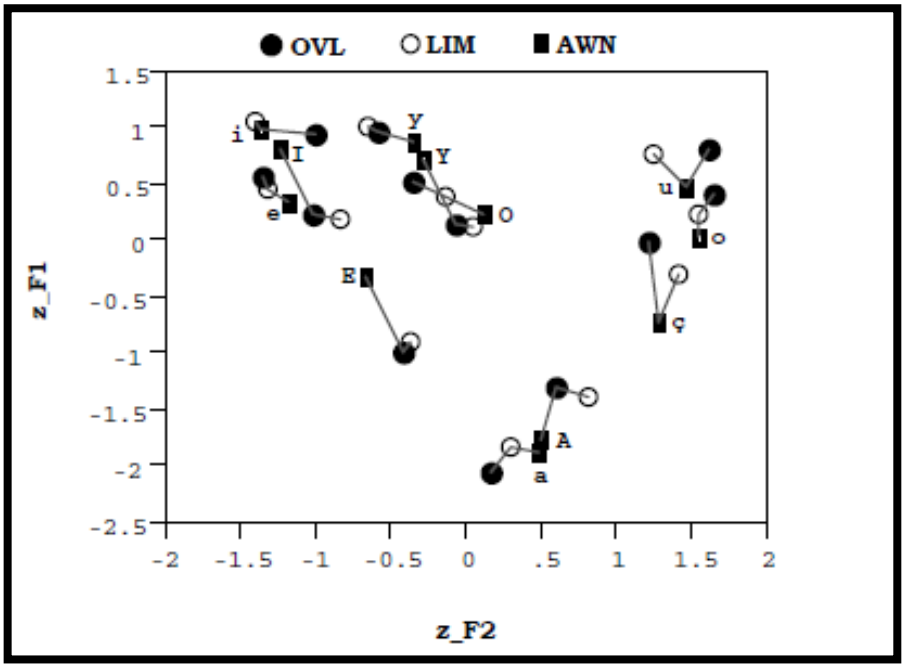


Figure 4.5. Average z-score [a vowel normalisation technique] for the Dutch monophthongs in male and female speakers from the Antwerp (AWN), Limburg (LIM) and East-Flemish (OVL) region (Verhoeven & Van Bael, 2002a: 157).

In the Brabantine dialect area as a whole, these quantitative rather than qualitative distinctions generally are a disappearing relict dialect feature, threatened by the emergence of the Standard Dutch realisation of /ɪ/ and /ʏ/; the Southern Brabantine area, nevertheless, is regarded as an exception, with the Antwerp city dialect being a northern offshoot that also holds on to this feature, as described in the *Dictionary of Brabantine dialects* (Belemans *et al*., 2000: 113). The Antwerp city dialect is an expanding variety (Taeldeman, 1999: 286) within "the economically and culturally dominant region in [the Flemish speech community]", as Adank *et al.* characterise the Brabantine region (2004a: 1730). It is thus expected that this durational distinction remains relatively unchallenged and is still present in the majority of speakers from the Antwerp area.

As already mentioned, Verhoeven & Van Bael (2002a&b) provide acoustic measurements that support the impressionistic observations in other works. They point out that "the vowel realisations in Antwerp Brabantine Dutch differ substantially from those in the other regions" (2002a: 161). The reading task the authors set up aimed to elicit (regional varieties of) Standard Dutch, and the participants were regarded as producing (regional varieties of) Standard Dutch. However, as can be seen in Figures 4.4 and 4.5 above, the (normalised) /i-ɪ/ and /y-ʏ/ realisations of Antwerp speakers are spectrally almost identical, and durationally relatively more distinct (in comparison with the other two varieties). This durational rather than spectral difference is remarkable, since Verhoeven & Van Bael (2002a&b) tested not for dialectal Dutch productions, but for regional varieties of Standard Dutch, with standardised carrier sentences containing the target vowels. This provides evidence for the impressionistic observation made in Belemans *et al.* (2000) about the 'problematic' production of Standard Dutch /ɪ/. This is also echoed in Ooms & Van Keymeulen, where the authors note that the durational instead of the spectral contrast is deeply rooted, and that Antwerp speakers moreover do not realise that their contrast is

deviating, causing it to be transferred even when attempting to speak Standard Dutch. (2005:30).

It should be noted that this last publication is a popularising work rather than a scientific study. However, together with the other data and observations provided here, it sketches an image of Antwerp (Brabantine) Dutch that is not at all affected by the spectral contrast present in Standard Dutch. The theoretical models discussed in section 2 attach a great deal of importance to the L1 production environment to account for L2 perception and subsequent production. If the durational distinction in the two vowel contrasts is an intricate part of the L1 production environment of people from Antwerp, and is, as discussed, not influenced by Standard Dutch, this could very well mean that the Antwerp participants perceive (and produce) L2 vowels differently than speakers of regional varieties that are much more similar to Standard Dutch in the relevant vowel contrasts, as is the case for Ghent speakers.

4.2 L2 vowel system

4.2.1 RP English

Collins & Mees classify the RP English /i:/ vowel as a "[c]entred from front close, unrounded, free, steady-state vowel" (2003: 100); the /ɪ/ vowel, on the other hand, is classified as a checked vowel (2003: 90). Collins & Mees further classify checked /ɪ/ as a "[f]ront-central, close-mid, unrounded, steady-state vowel" (2003: 90). In comparison with checked vowels, English free vowels are longer in similar phonetic contexts (Collins & Mees, 2003: 100). There is thus a durational as well as a spectral distinction in the English /i:-ɪ/ contrast. Hawkins & Midgley (2005), who measured the formant frequencies of RP vowels in four age groups, provide evidence for the spectral distinction in this contrast, as can be seen in Figure 4.6 below (in which results from Wells, 1962 and Deterding, 1990 are also included).

Collins & Mees point out that speakers from Antwerp may indeed experience problems with English /i:-ɪ/ contrast:

> *"Note that certain speakers of urban accents, e.g. The Hague, Rotterdam, Amsterdam, Antwerp, may have a [Dutch] /ɪ/ which is closer and more front, sounding similar to types of [English] /i:/. These may sound dialectal if used for [English] /ɪ/, or cause confusion between [English] /ɪ - i:/"* (2003: 91).

The results from Debaene (2012) suggested that there was indeed an overlap in the English /i:/ and /ɪ/ productions by Antwerp speakers.

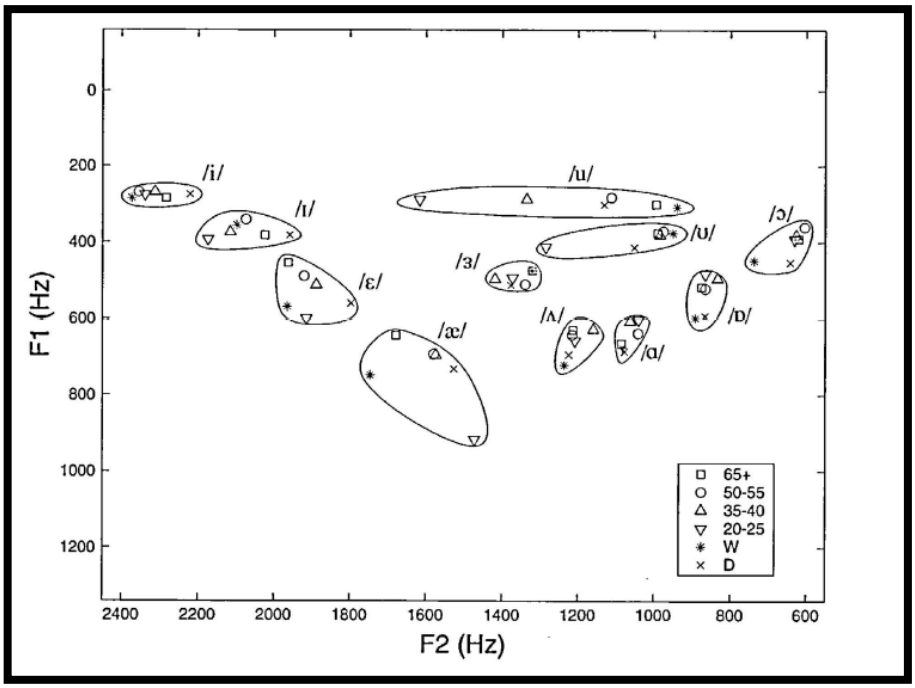


Figure 4.6. Mean frequencies (Hz) of the first and second formants of each of the eleven RP monophthongs, for each age group of Hawkins & Midgley (2005), together with those of Wells (1962) ('W') and Deterding (1990) ('D'). The lines surrounding each vowel phoneme separate the vowel distributions but have no statistical status. (taken from Hawkins & Midgley, 2005: 186).

Dutch /y/ does not have an English counterpart that is transcribed with the same IPA character, nor does Dutch /ʏ/. It was therefore decided to investigate the perception and production of the English vowels /ʊ/ and /ʌ/. Collins & Mees characterise /ʊ/ as "[b]ack-central, close-mid, slightly rounded, checked, steady-state", and /ʌ/ as "[c]entral-front, below open-mid, unrounded, checked, steady-state" (2003: 94). These two vowels can thus not be seen as equivalent to the /y/ and /ʏ/ vowels, yet there are several reasons to investigate the perception and production of these two vowels. First, although there is no durational difference between them, they still are spectrally distinct from each other. Secondly, as noted in Harrington, Kleber & Reubold (2011), and as can also be seen in Figure 4.6, /ʊ/ tends to shift towards a more fronted realisation, causing it to be spectrally more similar to Dutch /y/. Concerning the /ʌ/ vowel, Collins & Mees (2003: 95) point out that this vowel is often substituted for Dutch /ʏ/ (they transcribe it as /ʉ̟/), so it seems interesting to investigate whether the Dutch regional differences in the /y-ʏ/ contrast influence the perception and production of this vowel.

4.3 Overview

To conclude the discussion on the L1 standard and dialectal vowel systems, Table 4.2 provides an overview of the vowel pairs discussed in the present section. It should be noted that the classifications regarding spectral and durational contrasts between the vowels are simplifications used for methodology's sake.

Table 4.2. Overview of the relevant vowels/vowel pairs from the vowel systems under consideration in section 4 [11].

Variety	Transcription	Durational contrast	Spectral contrast	References
Southern Standard Dutch	/i-ɪ/	limited (some studies: /i/ = half-long)	Yes	Nooteboom (1971); Koopmans - van Beinum (1980); Van Bael & Verhoeven (2002a&b); Adank, van Hout & Smits (2004); Rietveld et al. (2004); Leussen, Williams & Escudero (2011)
	/y-ʏ/	limited (some studies: /y/ = half-long)	Yes	
Ghent Dutch	/i-ɪ/	limited	Yes	Taeldeman (1985); Van Bael & Verhoeven (2002a&b)
	/y-ʏ/	limited	Yes	
Antwerp Dutch	/iː-ɪ/	Yes	limited	Nuyts (1989); Belemans et al. (2000); Van Bael & Verhoeven (2002a&b); (Ooms & Van Keymeulen (2005))
	/yː-ʏ/	Yes	limited	
English RP	/iː - ɪ/	Yes	Yes	Collins & Mees (2003); Hawkins & Midgley (2005); Harrington, Kleber & Reubold (2011)
	/ʊ/ & /ʌ/	No (note: no vowel 'pair')	Yes (note: no vowel 'pair')	

[11] Because Antwerp /iː yː/ and Ghent and Standard Dutch /i y/ are regional allophones of the same phoneme, and so as to avoid confusion with English /iː/, the Antwerp variant of Dutch /i/ will be transcribed as /i/ in the present study.

5 Perception experiment

5.1 Aim

The aim of the perception experiment, which consists of two different tasks, is twofold. A Dutch perceptual categorisation task first examines whether there are regional differences between Antwerp and Ghent listeners in the perception of Dutch vowel pairs /i-ɪ/ and /y-ʏ/ produced by a Standard Dutch speaker. This is done by examining onto which native phonological categories these vowels are mapped by listeners from the two different dialect areas.

In a second, cross-language (English-Dutch) perceptual categorisation task, listeners perceive the RP English vowels /iː ɪ ʊ ʌ/, but they are led to believe that the perceived stimuli are still Dutch sounds. It is then examined whether there are differences between the two dialectally differing participant groups with respect to the L1 Dutch vowel categories onto which the RP English vowels are mapped, and whether these potential differences can be explained by analysing them in the light of the Dutch perception task results.

5.2 Methodology

5.2.1 Participants

This experiment examined the Dutch and English vowel perception of 12 Dutch listeners from the Ghent area (four of which were male) and 12 Dutch listeners from the Antwerp area (two of which were male). This majority of female participants should not be ignored, as it could have an influence on the results: it is believed that female speakers generally tend to use standard language variant more than male speakers (Mesthrie *et al.*, 2000: 102).

All participants completed a questionnaire about their basic personal information and language background. The Ghent area participants were aged between 19 and 23 years at the time of testing (with one exception of a participant who was aged 26). The participants from the Antwerp area were aged between 18 and 21 years at the time of testing. All participants were higher education students. The native language of all participants was Dutch, as was the language spoken at home (two Ghent participants also spoke French at home). The mother tongue of the participants' parents was Dutch as well, except for the father of one Ghent and one Antwerp participant (who had Lingala and French as a mother tongue, respectively). Apart from French and German, which are taught to most students in Flemish secondary schools, some participants also knew Spanish and/or Italian.

Like the majority of Flemish people in that age group, none of the participants had received English instruction prior to the age of twelve years, and none of them had stayed in an English-speaking country for a longer period of time (except for a Ghent participant who stayed in England for three months in the context of the Erasmus Programme at the age of 21). In accordance to, among others, Flege 1995 (cf. section 2) and Newport 1990, who propose a linear relationship between the age of L2 learning and L2 proficiency, they can thus be considered to take an intermediate position on the 'younger learner vs. older

learner' scale. With Flege's SLM (1995) and Escudero's L2LP (2005) in mind, native-like perception and production (though harder than for early learners) could still be possible for the participants of the present study.

Concerning English education, the participants had received 5 to 6 years of English classes during their formal secondary education, ranging from 2 to 4 hours per week. Furthermore, it was ascertained that none of the participants (had) studied English or Dutch at university level, in order to rule out plausible influence of intensive Standard Dutch or Standard English (RP-based) pronunciation training, and to prevent influence of academic knowledge of (and thus disadvantageous awareness of) the characteristics of the own regional L1 variety.

Participants were paid 5 euros for their participation.

The answers to the questionnaire are presented in their entirety in Appendix A.

5.2.2 Design of the task and stimuli

The perception experiment consists of two parts: in the first part, listeners from the two different regional backgrounds perceptually map Standard Dutch stimuli onto Standard Dutch phonological categories through a categorisation task. In the second part of the task, which also consists of a categorisation task, the participants hear RP stimuli, but they are led to believe that these stimuli are Dutch stimuli as well. This way, it was possible to investigate how non-native vowels are mapped on the L1 categories, and if there were any differences between the two participant groups from different regional backgrounds. It was decided that the perception experiment in this case study would be conducted before the production experiment. This was done mainly because of the fact that the production experiment also contains a two-fold task, but with a distinct Dutch and English part, which was also communicated as such (cf. section 6.2.2). It was likely that the participants would infer from the bilingual character of the production experiment that the perception

experiment would also be bilingual (which is the case, but which was not explicitly revealed to the participants).

The stimuli for the Dutch task were produced by a young female native speaker of Standard Dutch who teaches Standard Dutch at university level. In analogy, the stimuli for the English task were produced by a young female native speaker of Southern British English who teaches RP pronunciation at university level. To rule out influences of different consonantal contexts and knowledge of existing words, it was decided that the stimuli for the perception experiment would be pronounced in an identical consonantal environment, to create non-words. The specific consonantal context opted for was /h V s/, because noteworthy consonantal influence was not expected in this environment; furthermore, this consonantal context rendered mainly non-words: only /h ʊ s/ could be interpreted as the Dutch word 'hoes', which means 'cover' or 'sleeve'.[12] Moreover, the consonantal context occurs both in Dutch and in English, minimising even further the influence of the consonantal context. The stimuli obtained in this way are presented in Table 5.1.

Table 5.1. The Dutch and English stimuli included in the Dutch and the English categorisation tasks.

Dutch	
vowel	stimulus
/i/	/h i s/
/ɪ/	/h ɪ s/
/y/	/h y s/
/ʏ/	/h ʏ s/
English	
vowel	stimulus
/iː/	/h iː s/
/ɪ/	/h ɪ s/
/ʊ/	/h ʊ s/
/ʌ/	/h ʌ s/

The Dutch stimuli were repeated 4 times each by the native Standard Dutch speaker, and the same was done with the English stimuli by the RP English speaker, providing a total

[12] It could be argued that /h ɪ s/ renders the English word 'hiss'; however, the assumption of the listeners that the second, cross-linguistic task also contained Dutch stimuli is expected to be kept to a minimum.

of 16 unique stimuli per task. Appendix B presents the averages, standard deviations and means of the spectral and durational features of the target vowels produced by the two native speakers. As can be derived from the standard deviations, the stimuli did not differ much from each other (due to the identical consonantal environment), which made them ideally suited for this task.

In the task itself, which was designed with the use of Praat (version 5.3.45, Boersma & Weenink, 2013), these stimuli were played twice, providing a total of 32 categorised stimuli per participant per task. A short pause halfway through was programmed into the task, as well as a command that randomised the order in which the stimuli were played. During the task, the participants were presented with a grid that contained 12 Dutch monophthongs. These vowels were presented in orthographic representations of common monosyllabic words (cf. Table 5.2). This orthographic representation was opted for since the participants were not expected to be familiar with IPA transcriptions; moreover, there are no regional differences in orthography: the spelling of Dutch words is the same in Antwerp as in Ghent.

Table 5.2. Targets for the two categorisation tasks with their phonological transcriptions and translations.

kus /k ʏ s/ 'kiss'	muur /m y r/ 'wall'	deur /d ø: r/ 'door'
zeep /z e: p/ 'soap'	kaas /k a: s/ 'cheese'	kop /k ɔ p/ 'cup/head'
hoop /h o: p/ 'hope/pile'	boek /b u k/ 'book'	dief /d i f/ 'thief'
pit /p ɪ t/ 'kernel'	bed /b ɛ t/ 'bed'	dak /d ɑ k/ 'roof'

5.2.3 Procedure

The Standard Dutch stimuli for the Dutch task were recorded by the author with a Marantz PMD620 digital recorder and a Sony ECM-MS907 one-point stereo microphone, which were also used in recording the production tasks. Since the native English speaker was not readily available in person, she recorded the stimuli herself and sent them via e-mail

attachment. These English stimuli were recorded with an Olympus VN5500PC Digital Voice Recorder (with built-in microphone).

The perception experiment was conducted in the same session as the production experiment, which followed immediately after the perception tasks. The participants were briefly explained how the task was to be conducted. A printed screenshot (cf. Figure 5.1) was shown on beforehand to ensure familiarity with the task design. They were told not to pay attention to the productions as such, as these were all non-words: it was the vowel that needed to be focused on. To avoid confusion, it was explained that due to the randomisation of the stimuli, it was perfectly possible that two or more similar-sounding sounds could follow one another, and that the only thing which was important was the sound that they thought they heard. As for the second (English) perception task, they were told that this was again a Dutch categorisation task, but with another speaker now.

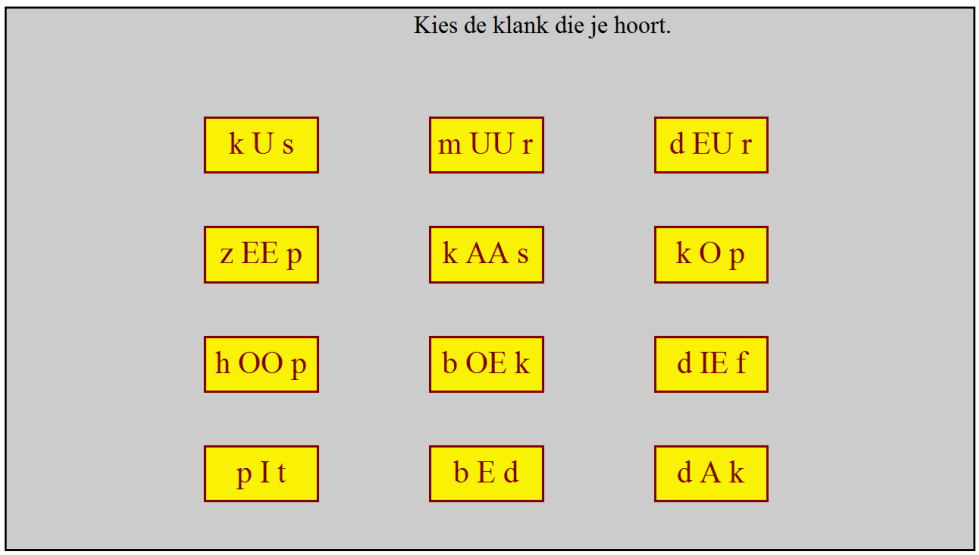


Figure 5.1. Screenshot of the task lay-out. Participants were instructed to click on the sound which they thought they heard ('Kies de klank die je hoort' translates to 'Choose the sound that you hear').

5.2.4 Coding and analysis

First, the stimuli produced by the native Standard Dutch and RP English speakers were measured to facilitate a cross-language comparison of the participants' results (the results of these measurements are presented in Appendix B). The vowel duration and the first two

formant frequencies were measured with Praat (version 5.3.45, Boersma & Weenink, 2013). Although there are more than two formants, "[i]t is generally assumed that F1 and F2 are most relevant for a vowel's identity" (Verhoeven & Van Bael, 2002a: 149). Hence, only these two formants are taken into consideration in the measurements of the target stimuli in the perception experiment (and also of the productions in the production experiment - cf. section 6.2.4). While acknowledging the complex connection between the acoustic realisations of vowels and their articulation, Verhoeven & Van Bael report that "it is accepted that F1 mainly correlates with articulatory degree of opening, while F2 reflects place of articulation" (2002a: 149). A higher F1 frequency corresponds to a lower vowel height, while a higher F2 frequency corresponds to a greater frontedness in the production of a vowel (cf. Figure 5.2).

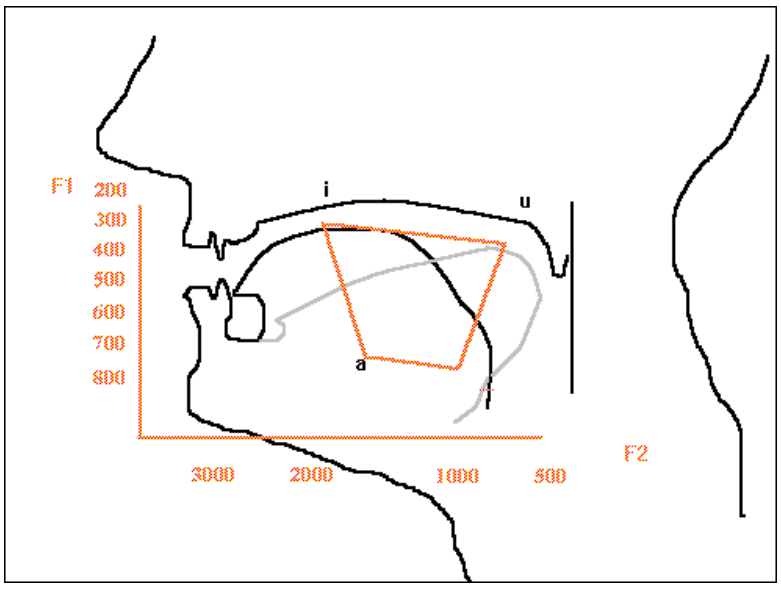


Figure 5.2. Visual representation of how the height and frontedness/backness of the tongue body relate to the respective F1 and F2 values (taken from Gramley ,2010).

After each perception task, the results were saved in text files, which presented for each of the 32 stimuli the type of stimulus, and which of the 12 categorisation possibilities the participant had chosen. Each of the four English and Dutch target vowels yielded 96 categorisations per participant group, for a total of 768 categorisations per task. The results, calculated per participant group and per task, are presented in section 5.4.

To enable a thorough comparison of the Ghent and Antwerp listener data with regard to the particular target stimuli from the Dutch or English task, a statistical analysis was performed. For the present experiment, the statistical tool SPSS (IBM Corp., 2012, version 21.0) was used. To enable a processing of the data, the twelve target vowels are seen as twelve sound categories, each of which was assigned a numerical value: a value of 1 corresponds to the /i/ category; a value of 2 corresponds with the /ɪ/ category, etc. For each Dutch and English target stimulus, and per dialect, the respective 96 observations are labelled according to which of the twelve sound categories it was mapped onto. It should be noted that no specific arrangement or hierarchical ordering between categories is implied in the use of increasing numerical values: this methodology is only used to distinguish between the different vowel categories in some way, and to make statistical calculations possible. In the present experiment, sound is a nominal variable with twelve categories in which the sounds are classified.

The statistical test used for this experiment is a (two-tailed) Pearson chi-squared test - the best-known statistical hypothesis tests in linguistics, according to Triest (2009b) - which is applied to two independent samples of nominal variables (in this case the Ghent and Antwerp vowel categorisations). This test presupposes that the samples used are independent, which is the case for the present experiment: although certain criteria for participant selection had to be kept in mind (cf. 5.1.3.1), the selection of Ghent participants did not depend on the selection of Antwerp participants, or vice versa.

For each target stimulus (/his hɪs hys hʏs/ for Dutch, /hiːs hɪs hʊs hʌs/ for English) there are 192 categorisation observations (96 per participant group). Each of the observations is assigned a numerical value ranging from 1 to 12, in accordance with which of the twelve Dutch target vowels it was mapped onto. This is the first nominal variable (target vowel category); the other nominal variable is region, and has a value of 1 for Antwerp and 2 for Ghent. Grouping the first nominal variable (vowel categorisations)

according to the second nominal variable (region) yields two samples per stimulus (one Ghent and one Antwerp sample). For these two variables (target vowel category and region), a contingency table is composed, an example of which is provided in Table 5.3.

Table 5.3. Contingency table for the /hɪs/ stimulus from the Dutch perception task.

Region	CATEGORY												total
	1 /i/	2 /ɪ/	3 /eː/	4 /ɛ/	5 /aː/	6 /ɑ/	7 /oː/	8 /ɔ/	9 /øː/	10 /u/	11 /y/	12 /ʏ/	
Antwerp	6	86	0	4	0	0	0	0	0	0	0	0	96
Ghent	0	95	0	0	0	0	0	0	0	0	0	1	96
total	6	181	0	4	0	0	0	0	0	0	0	1	192

For the implementation of the Pearson chi-squared test, the expected counts for each of the so-called 'cells' of this table (12 Antwerp and 12 Ghent cells in which the stimuli can be categorised) are calculated on the basis of the contingency table from Table 5.3. The expected counts per cell are calculated as ('total of the row' multiplied by 'total of the column') divided by ('stimuli total').[13] Table 5.4 presents the contingency table from Table 5.4, now with the expected counts inserted for each cell.

Table 5.4. The expected counts for the 12 Antwerp and Ghent cells for the /hɪs/ stimulus from the Dutch perception task.

Region	CATEGORY												total
	1 /i/	2 /ɪ/	3 /eː/	4 /ɛ/	5 /aː/	6 /ɑ/	7 /oː/	8 /ɔ/	9 /øː/	10 /u/	11 /y/	12 /ʏ/	
Antwerp	3	90,5	0	2	0	0	0	0	0	0	0	0,5	96
Ghent	3	90,5	0	2	0	0	0	0	0	0	0	0,5	96
total	6	181	0	4	0	0	0	0	0	0	0	1	192

[13] The expected count for the Antwerp category 1 (/i/) cell, for instance, is calculated as follows: ((6+86+4) x (6+0))/ (192) = 3.

For the categories indicated in red in Table 5.4, the amount of observations - and subsequently the expected counts - is zero. Therefore, the Pearson chi-squared test does not take the values for these categories into account in its calculations, as they do not affect the outcome of the test.

Two conditions for the correct implementation of the Pearson chi-squared test are that no more than 20% of the cells can have an expected count of less than 5, and none of the cells can have an expected count of less than 1 (De Pelsmacker & Van Kenhove, 2012: 243).[14] Table 5.4 shows that six of the remaining eight cells (66.67%) have an expected count of less than 5 (indicated in yellow). This is more than the prescribed maximum of 20%. Moreover, the expected count is 0.5 for category 12, which is lower than 1, thus violating the second condition. When repeating this procedure for each of the eight stimuli, it becomes apparent that this transgression is no exception: seven of the eight stimuli reveal a distribution of the data that would lead to a distorted result of the Pearson chi-squared test.

Following De Pelsmacker & Van Kenhove (2012: 243), a (two-tailed) Fisher's exact test is alternatively used in those cases where the expected counts are too low. In the present experiment, the Pearson chi-squared test could only be used for English /hʌs/; for the other stimuli, Fisher's exact test had to be used. The Pearson chi-squared test and Fisher's exact test both have the following null and alternative hypotheses:

- Null hypothesis (H_0): the distribution into different sound categories of a certain Dutch or English stimulus is identical between Ghent and Antwerp listeners;

[14] It should be noted that the cells with an expected count of 0 are not taken into account for this condition, as they are not taken into account in the test itself (cf. previous paragraph).

- Alternative hypothesis (H_A): the distribution into different sound categories of a certain Dutch or English stimulus is not equal between Ghent and Antwerp listeners.

This way, the test verifies whether or not the Ghent and Antwerp categorisation distributions are significantly different. A p-value of less than 0.05 - the norm for social sciences, according to Triest (2009b) - indicates that the null hypothesis is to be rejected and that there is a significant difference between, in this case, Ghent and Antwerp distributions on a significance level of 5%.

Figure 5.3 presents an example of the SPSS output of the statistical processing, again for the Dutch /hɪs/ stimulus. Sentence a in the output indicates which test is used: if less than 20% of the cells have an expected count of less than five, the p-value from the Pearson chi-squared test is used ('Asymp. Sig. (2-sided)'). If more than 20% of the cells have an expected count of less than five (cf. Figure 5.3), the p-value from Fisher's exact test ('Exact Sig. (2-sided)') is used. The p-value of 0.002 in Figure 5.3 indicates a significant regional difference in the categorisation of this particular stimulus at a 5% significance level.

hisDuCategorie * regio Crosstabulation

			regio		
			1,00	2,00	Total
hisDuCategorie	1,00	Count	6	0	6
		% within regio	6,3%	0,0%	3,1%
	2,00	Count	86	95	181
		% within regio	89,6%	99,0%	94,3%
	4,00	Count	4	0	4
		% within regio	4,2%	0,0%	2,1%
	12,00	Count	0	1	1
		% within regio	0,0%	1,0%	0,5%
Total		Count	96	96	192
		% within regio	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	11,448[a]	3	,010	,002		
Likelihood Ratio	15,697	3	,001	,002		
Fisher's Exact Test	11,162			,002		
Linear-by-Linear Association	,525[b]	1	,469	,732	,366	,076
N of Valid Cases	192					

a. 6 cells (75,0%) have expected count less than 5. The minimum expected count is ,50.

b. The standardized statistic is ,725.

Figure 5.3. SPSS output for the /hɪs/ stimulus from the Dutch perception task (the appropriate test and its p-value are indicated in green).

5.3 Hypotheses

The results from previous studies presented in section 3 and the differences between the regional Dutch vowel systems discussed in section 4 suggest that regional L1 influence on the perception of RP English vowels is likely to take place. In the present subsection, hypotheses are forwarded about how these presumed regional differences could influence performance in the Dutch and cross-language perception tasks. Since the vowel distribution and differentiation within Ghent (East-Flemish) Dutch resembles that of Standard Dutch, it is predicted that on the whole, these listeners will be more successful in mapping the vowels according to how they are intended to be mapped in comparison with the Antwerp listeners.

5.3.1 Dutch perception task: /i/ and /ɪ/

For this contrast, it is hypothesised that Antwerp listeners will have more difficulties than Ghent listeners. The Ghent (East-Flemish) Dutch /i-ɪ/ contrast bears a spectral and durational resemblance to Standard Dutch /i-ɪ/. The hypothesis is thus that the Ghent listeners will accurately categorise this contrast - in Flege's SLM (1995) terms, the phonetic categories Ghent listeners have constructed (based on their regional production environment) resemble those of Standard Dutch.

The phonetic categories constructed by Antwerp listeners, on the other hand, are weighted towards a durational rather than a spectral distinction. Therefore, Antwerp listeners will have more difficulties with Standard Dutch /i/ and /ɪ/ (distinguished spectrally rather than durationally).

Standard Dutch short /i/ could be regarded, in terms of Flege's SLM (1995), as *similar* both to Antwerp /ɪ/ (durationally and spectrally) and to Antwerp /i/ (spectrally). Comparatively, Standard Dutch /i/ could therefore be categorised more often as /ɪ/ by Antwerp listeners than by Ghent listeners. Since Antwerp /i-ɪ/ coincide spectrally, Antwerp

listeners could map spectrally deviating Standard Dutch /ɪ/ either onto /i/ or /ɪ/.[15] However, durational differences make it seem less likely that Standard Dutch /ɪ/ will be mapped onto /i/ ([i:] in Antwerp (Brabantine) Dutch) than vice versa, and Antwerp listeners are thus expected to have more difficulties with Standard Dutch /i/ than with Standard Dutch /ɪ/.

5.3.2 Dutch perception task: /y/ and /ʏ/

The hypothesis here is that the Ghent listeners will be more accurate than the Antwerp participants in their /y-ʏ/ mapping, as the Ghent /y-ʏ/ contrast is more similar to the Standard Dutch one than to the Antwerp one.

The Antwerp /y-ʏ/ contrast, on the other hand, bears similarities to the Antwerp /i-ɪ/ contrast: there is again a distinction on the basis of durational rather than spectral qualities (cf. Figure 4.5 in section 4.1.3). Consequently, correct categorisation of Standard Dutch /y-ʏ/ by the Antwerp listeners is likely to depend on durational rather than spectral features.

For Antwerp listeners, a similar prediction can be made for Standard Dutch /ʏ/ as for Standard Dutch /ɪ/. The spectral qualities of Standard Dutch /ʏ/ differ from those of spectrally coinciding Antwerp /ʏ/ and /y/, but durational differences make it seem less likely that Standard Dutch /ʏ/ will be mapped onto /y/ ([y:] in Antwerp (Brabantine) Dutch) than vice versa. Standard Dutch /y/ is thus expected to be mapped incorrectly by Antwerp listeners more often than Standard Dutch /ʏ/.

[15] Antwerp /ɛ/, as Figure 4.5 in section 4.1.3 shows, is spectrally raised so that it impressionistically could be perceived as Standard Dutch /ɪ/. However, Debaene (2012) (in which Antwerp /ɛ/ was researched) found no influence of this deviant spectral feature on RP English /ɪ/ production. Given both Escudero's assertion that production follows perception (2005: 326) (and could thus be regarded as a retroactive indicator for perception), and the fact that English and Standard Dutch /ɪ/ are, according to Flege's (1995) terminology, *identical* (or at least *similar*), it is improbable (though not impossible) that Antwerp listeners categorise the Standard Dutch /ɪ/ stimuli as /ɛ/.

Figure 4.5 shows that the Limburg and East-Flemish (and thus also the expected Standard Dutch) realisations of /ʏ/ lie spectrally close to the long Antwerp /ø:/ vowel (transcribed as 'O' in the figure). However, Figure 4.3 in section 4.1.2 shows that /ø:/ produced by Antwerp (as well as Limburg and East-Flemish) speakers is approximately twice as long as Limburg and East-Flemish /ʏ/ productions, no mapping onto this phonological category is expected.

As the participants were led to believe that the RP stimuli they heard in the second (English) perception task were Dutch stimuli, sections 5.3.3 and 5.3.4 will look into how the mapping of these English vowels onto Dutch vowel categories could be influenced by the English vowel properties, and whether regional differences can be discerned in this perceptual mapping.

5.3.3 Cross-language perception task: /i:/ and /ɪ/

The RP English /i:-ɪ/ vowel pair has a two-dimensional (spectral and durational) distinction, corresponding in part to the spectral Ghent (East-Flemish) Dutch distinction, and in part to the durational Antwerp (Brabantine) Dutch distinction. The hypothesis is that the additional durational distinction in RP English (which is not present in Standard Dutch) will cause Antwerp listeners to categorise RP English /i:/ as /i/ more often than Standard Dutch /i/. In Flege's SLM (1995) terms, RP English /i:/ is *identical* to Antwerp (Brabantine) Dutch /i/, as both vowels are long ([i:]) and produced with similar spectral qualities. RP English /ɪ/ could then be *similar* to Antwerp /ɪ/, causing it to be categorised as such through equivalence classification (Flege, 1995)/perceptual assimilation (Best & Tyler, 2007).

This /i:-ɪ/ contrast is not expected to be problematic for the Ghent listeners. The Ghent spectral distinction is produced in RP English, too; an additional durational difference only contributes to a greater distinction between the two vowels, and to an even

more "successful" categorical mapping onto the L1 Dutch categories than for Standard Dutch.

The predictions for Antwerp listeners show that an L1 regional vowel contrast that is more similar to an L2 contrast than to an L1 standard language contrast could lead to a situation in which the listeners from this region perform more accurately in perceiving this L2 contrast than in perceiving the L1 standard language contrast. A different L2 contrast (i.c. durational and spectral in RP English, instead of only spectral in Standard Dutch) does not always lead to incorrect perceptual mappings: if the vowel distinction is more pronounced in the L2 than in the L1, or if it resembles a regional distinction in some way (i.c. the durational Antwerp /i-ɪ/ contrast), more accurate L2 performance can be possible.

5.3.4 Cross-language perception task: /ʊ/ and /ʌ/

These vowels are investigated because their perceptual mapping by Ghent and Antwerp listeners may be influenced by regional differences in the /y-ʏ/ contrast.

Dutch only has short /u/ in the spectral vowel space in which RP English has both short /ʊ/ and long /u:/. Consequently, it is hypothesised that RP English /ʊ/ will mostly be perceptually assimilated to Dutch /u/. However, English /ʊ/ differs spectrally from Dutch /u/ (following Flege's 1997 classification, a different IPA symbol even suggests a classification of this sound as *new*). /ʊ/ is produced lower and more front[16], so that it could be perceived (to a lesser extent) as more fronted vowels, such as /y/ and /ʏ/.

/ʊ/ is predicted to be mapped more often onto /ʏ/ by Antwerp listeners than by Ghent listeners. This is mainly because RP English (short) /ʊ/ could be perceived as Antwerp (short) /ʏ/, but it would be too closed to be perceived as Ghent (short) /ʏ/. Nevertheless, it remains possible that /ʊ/ is perceived as (rather short) Ghent /y/ (but not

[16] This is especially the case with the recent phenomenon of the fronting of high back vowels in young Southern British English speakers (e.g. Hawkins & Midgley, 2005; Chládková & Hamann, 2011 for /u/)

as much as Antwerp longer /y/). Perceptual assimilation to Dutch /ø:/ cannot be ruled out either (on a spectral level), yet this is not likely due to /ø:/ being durationally long.

According to Collins & Mees (2003: 95), RP English unrounded /ʌ/ tends to be substituted with Dutch rounded /ʏ/ by Dutch speakers, even though Dutch /ʏ/ does not spectrally resemble English /ʌ/.[17] It is hypothesised that, upon only hearing English /ʌ/, listeners will rather categorise it as Dutch /a:/ or /ɑ/, two vowels which resemble English /ʌ/ in that they are also unrounded and lie in the same spectral vowel space.

The hypothesis is that perceptual assimilation of /ʌ/ to Dutch /ʏ/ will occur more for Ghent listeners, as Ghent /ʏ/ lies spectrally closer to RP /ʌ/ than Antwerp /ʏ/. Figures 4.3 and 4.5 moreover show another discrepancy between Ghent (East-Flemish) and Antwerp (Brabantine) Dutch: Antwerp /ɑ/ (as opposed to East-Flemish /ɑ/) lies spectrally close to Antwerp /a:/ (and thus to English /ʌ/), while being roughly twice as short. This could lead Antwerp listeners to categorise short RP /ʌ/ more often as /ɑ/, compared to Ghent listeners.

[17] Collins & Mees' prediction for this substitution is likely to be based in part on orthography: both English /ʌ/ and Dutch /ʏ/ are typically spelled as <u>.

5.4 Results

5.4.1 Dutch Perception Task

Tables 5.5 and 5.6 present the results of the Dutch perception task for Ghent and Antwerp listeners, respectively. The categorisations per participant are presented in Appendix C. Table 5.7 presents the p-values for the statistical tests performed on the results of the Dutch perception task. The corresponding SPSS outputs are presented in part 1 of Appendix D.

Table 5.5. Results of the Dutch perception task for Ghent listeners (note that because of the rounding off of the percentage values, these values do not always add up exactly to 100%).

Ghent		Target vowels											
		/i/	/ɪ/	/eː/	/ɛ/	/aː/	/ɑ/	/oː/	/ɔ/	/øː/	/u/	/y/	/ʏ/
Stimuli (total)	/h i s/ (96)	86 (89.6%)	7 (7.3%)	2 (2.1%)								1 (1.0%)	
	/h ɪ s/ (96)		95 (99.0%)										1 (1.0%)
	/h y s/ (96)		1 (1.0%)							5 (5.2%)		87 (90.6%)	3 (3.1%)
	/h ʏ s/ (96)		1 (1.0%)										95 (99.0%)

Table 5.6. Results of the Dutch perception task for Antwerp listeners (note that because of the rounding off of the percentage values, these values do not always add up exactly to 100%).

Antwerp		Target vowels											
		/i/	/ɪ/	/eː/	/ɛ/	/aː/	/ɑ/	/oː/	/ɔ/	/øː/	/u/	/y/	/ʏ/
Stimuli (total)	/h i s/ (96)	51 (53.1%)	44 (45.8%)	1 (1.0%)									
	/h ɪ s/ (96)	6 (6.3%)	86 (89.6%)		4 (4.2%)								
	/h y s/ (96)	1 (1.0%)	1 (1.0%)							14 (14.6%)		74 (77.1%)	6 (6.3%)
	/h ʏ s/ (96)									5 (5.2%)			91 (94.8%)

Table 5.7. The appropriate statistical test used per stimulus and the resulting p-value for the results of the Dutch perception task.

/h i s/	
test	p-value
Fisher's exact test	< 0.001
/h ɪ s/	
test	p-value
Fisher's exact test	0.002
/h y s/	
test	p-value
Fisher's exact test	0.066
/h ʏ s/	
test	p-value
Fisher's exact test	0.059

5.4.2 Cross-language Perception Task

Tables 5.8 and 5.9 present the results of the cross-language perception task for Ghent and Antwerp listeners, respectively. The categorisations per participant are presented in Appendix C. Table 5.10 presents the p-values for the statistical tests performed on the results of the cross-language perception task. The corresponding SPSS outputs are presented in part 2 of Appendix D.

Table 5.8. Results of the cross-language perception task for Ghent listeners (note that because of the rounding off of the percentage values, these values do not always add up exactly to 100%).

Ghent		Target vowels											
		/i/	/ɪ/	/eː/	/ɛ/	/aː/	/ɑ/	/oː/	/ɔ/	/øː/	/u/	/y/	/ʏ/
Stimuli (total)	/h iː s/ (96)	95 (99.0%)	1 (1.0%)										
	/h ɪ s/ (96)	1 (1.0%)	94 (97.9%)										1 (1.0%)
	/h ʊ s/ (96)										94 (97.9%)	2 (2.1%)	
	/h ʌ s/ (96)					42 (43.8%)	38 (39.6%)						16 (16.7%)

Table 5.9. Results of the cross-language perception task for Antwerp listeners (note that because of the rounding off of the percentage values, these values do not always add up exactly to 100%).

Antwerp		Target vowels											
		/i/	/ɪ/	/eː/	/ɛ/	/aː/	/ɑ/	/oː/	/ɔ/	/øː/	/u/	/y/	/ʏ/
Stimuli (total)	/h iː s/ (96)	82 (85.4%)	13 (13.5%)	1 (1.0%)									
	/h ɪ s/ (96)	1 (1.0%)	92 (95.8%)		3 (3.1%)								
	/h ʊ s/ (96)									1 (1.0%)	81 (84.4%)		14 (14.6%)
	/h ʌ s/ (96)					33 (34.4%)	55 (57.3%)						8 (8.3%)

Table 5.10. The appropriate statistical test used per stimulus and the resulting p-value for the results of the cross-language perception task.

/h iː s/	
test	p-value
Fisher's exact test	0.001
/h ɪ s/	
test	p-value
Fisher's exact test	0.246
/h ʊ s/	
test	p-value
Fisher's exact test	< 0.001
/h ʌ s/	
test	p-value
Pearson's chi-squared test	0.032

5.5 Discussion

5.5.1 Preliminary note: on the value of the statistical calculations

Before discussing the results, a few cautionary notes are in place. First, one should always keep in mind that the statistical calculations can display unreliable p-values in the present study; that a result is significant, does not necessarily mean that we can extend the observed differences to the whole population (or vice versa: that a result is not significant does not mean that there are no perceptual differences). The cause of this is the relatively limited sample of participants: with 12 Antwerp and 12 Ghent speakers, making generalisations should be done with caution. Furthermore, the participants were young, predominantly female higher education students - these participants are of course not necessarily representative for the whole Antwerp and Ghent population, and differences are likely to be more apparent in other age groups, in people with a lower education, in a predominantly male sample, etc. Because of this, population conclusions could lead to different p-values, and thus weaker or stronger conclusions.

It should also be pointed out that the participants performed the same task twice (i.e. an identical-looking task, in which only the stimuli differed). There is hence a chance that the answers on the English task are influenced - be it consciously or not - by the answers on the preceding Dutch task. Moreover, in the English task, the participants were already familiar with the design of the task, which could also be of some influence.

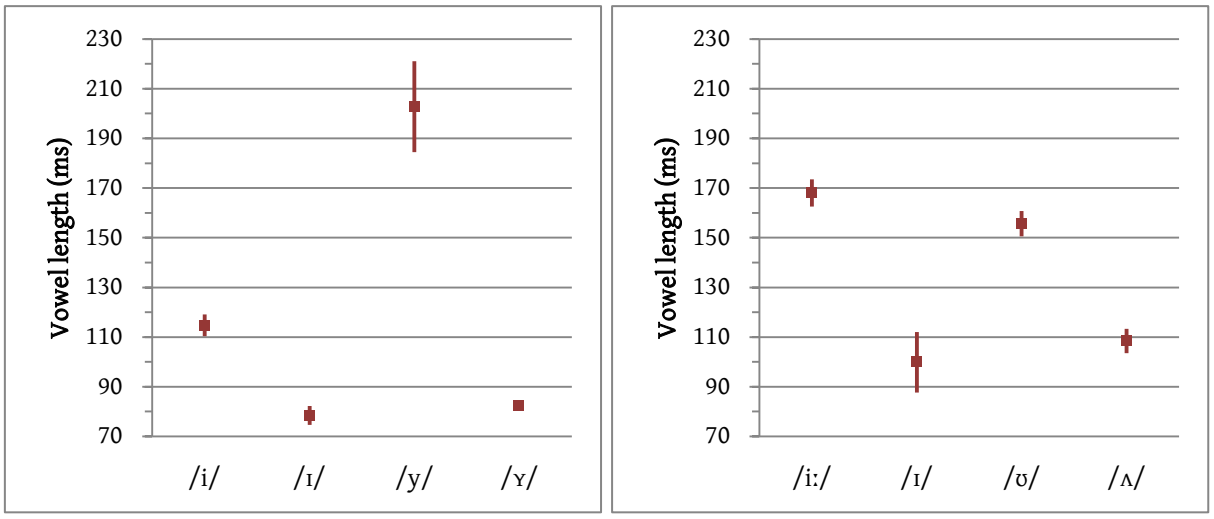
Secondly, the present study makes use of the terms 'Ghent listeners' and 'Antwerp listeners'. It should be noted that these terms refer to the respective Ghent and Antwerp samples from the present study, unless stated otherwise.

One last remark is that the statistical tests used only answer the question if significant cross-dialectal differences are apparent in vowel categorisation between Ghent and Antwerp listeners. The more the choice of categories or number of stimuli categorised

as a specific target vowel differ between the two listener groups, the more significant the test results will be. However, for the purpose of this study, it is also interesting to examine which vowel categories specifically are used to (wrongly) categorise the stimuli; even if this level of detail is not included in the statistical analysis, an assessment of the specific category choices can be indicative of cross-dialectal differences.

5.5.2 Preliminary note: graphic representation of the target stimuli

To facilitate the interpretation of the results presented in section 5.4, this subsection presents a graphic representation of the spectral and durational values of the target stimuli as produced by the native speakers (cf. Appendix B). Figures 5.3 and 5.4 show the average vowel duration of the target stimuli; Figure 5.5 shows the spectral features of the target stimuli.



Figures 5.3 and 5.4. Average vowel duration (in ms) of the Standard Dutch (left) and British RP English (right) target stimuli of the perception tasks. The vertical lines indicate the standard deviations.

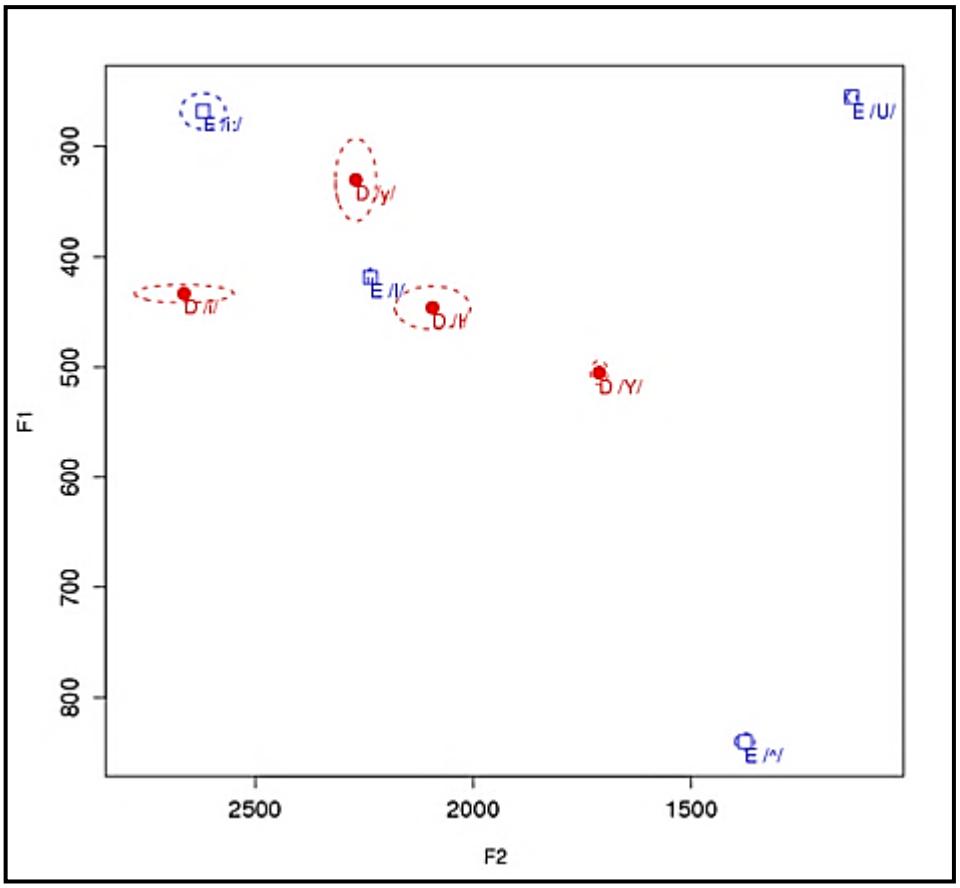


Figure 5.5. Mean values of the spectral features (F1 and F2 values, in Hz) of the Standard Dutch target stimuli (red) and the RP English target stimuli (blue). The ellipses indicate the standard deviations.[18]

5.5.3 Dutch perception task: /i/ and /ɪ/

The Ghent and Antwerp participants' results of the Dutch perception task for /ɪ/ and /i/ are consistent with the hypotheses proposed in section 5.3.1.

Concerning the perception of Standard Dutch /ɪ/, the twelve Ghent listeners categorised 95 of the 96 /hɪs/ stimuli accurately. This suggests that, according to Flege's SLM (1995), Ghent listeners have constructed a phonetic category for /ɪ/ that can also be applied to (or is based on) the perception of Standard Dutch /ɪ/, which is consistent with

[18] The NORM Vowel Normalization and Plotting Suite used to graphically represent the spectral features of the target stimuli does not support the use of IPA characters (Thomas & Kendall, 2007: http://ncslaap.lib.ncsu.edu/tools/norm/norm1_help.php). /ɪ/, /ʏ/, /ʊ/ and /ʌ/ are represented here as capital I, capital Y, capital U and '^', respectively.

the observation that the vowel qualities of Ghent and Standard Dutch are very similar (cf. section 4).

Although Table 5.7 shows a significant difference between the Ghent and Antwerp listeners for Dutch /ɪ/, the Antwerp listeners still show a high accuracy rate for the perception of Standard Dutch /ɪ/ (86 out of 96 stimuli, or 89.6%). In line with the smaller spectral differences between Antwerp /ɪ/ and /i/ than between Ghent /ɪ/ and /i/, 6 of the 96 /hɪs/ stimuli were categorised as /i/ by the Antwerp listeners (as opposed to none by the Ghent listeners). In agreement with the remarks about Antwerp /ɛ/ in section 5.3.1, 4 of the /hɪs/ stimuli were analysed as /ɛ/ (as opposed to none by Ghent listeners), suggesting subtle regional differences in the constructed phonetic categories for this vowel pair.

As for the perception of Standard Dutch /i/, the results are very outspoken. Ghent listeners categorise 89.6% of /his/ stimuli accurately. Antwerp listeners, however, only categorise 51 (or 53.1%) of the 96 /his/ stimuli accurately; 45.8% (44 out of 96) of the stimuli are categorised as /ɪ/. Table 5.7 reveals very significant differences between the two participant groups ($p < 0.001 < 0.05$, cf. Appendix D). While Ghent listeners are thus (as expected) comparatively accurate in perceiving Standard Dutch /i/, Antwerp listeners noticeably fail to consistently categorise this vowel accurately. The fact that Standard Dutch /i/ corresponds to Antwerp /i/ on a spectral level, and to Antwerp /ɪ/ on both a spectral and a durational level, could be the cause of this confusion and failure of correct categorisation. The results reveal that Antwerp listeners tend not to make a distinction between /i/ and /ɪ/ on a spectral level (or at least not on the basis of F2 or frontedness, as Figure 5.5 reveals that the F1 values of Standard Dutch /i/ and /ɪ/ are comparable[19]). Their phonetic categories for /i/ and /ɪ/ are not fine-tuned on the basis of spectral features, since

[19] Cf. Boersma & Chládková (2011) on how a view of cue constraints as referring to phonetically-based features explains accurate perception even when only one cue (e.g. or F1, or F2) is taken into account, and how this does not necessarily imply that only one cue is used in production as well.

there is no need to do so in the Antwerp dialect production environment. Standard Dutch /i/ could then be seen as *similar* to Antwerp /i/, but also to Antwerp /ɪ/, possessing characteristics of both these vowels. That slightly more /his/ stimuli were analysed accurately as /i/ than wrongly as /ɪ/ (53.1% vs. 45.8%), could in part also have to do with the target /i/ stimuli being durationally longer than the target /ɪ/ stimuli; a difference which is, however, not as outspoken as for Dutch /y-ʏ/ or English /iː-ɪ/ (cf. Figures 5.3 and 5.4). In all, correct perceptual assimilation of Standard Dutch /i/ does not seem to have taken place for Antwerp listeners.

It is not surprising that the accuracy rate of Antwerp listeners in categorising Standard Dutch /ɪ/ is much higher, since Standard Dutch /ɪ/ is durationally short, like Antwerp /ɪ/ (but unlike Antwerp /i/). Durational information thus seems to play a role in the accurate categorisation of Standard Dutch /ɪ/ by Antwerp listeners. If this were not the case, Antwerp listeners would be expected to categorise Standard Dutch /ɪ/ equally as /ɪ/ and as /i/. These two vowels, after all, are spectrally close(r) to each other for Antwerp listeners (and thus equally distant from Standard Dutch /ɪ/).

Furthermore, Standard Dutch /ɪ/ could be considered as *new* to Antwerp listeners. Indeed, no spectral counterparts of this vowel are present in the Antwerp dialect, and familiarity with Standard Dutch sounds (which is to be expected) would rather lead to a correct categorisation of Standard Dutch /ɪ/.

5.5.4 Dutch perception task: /y/ and /ʏ/

As hypothesised, Ghent listeners again performed more accurately in categorising Standard Dutch /y/ and /ʏ/ than Antwerp listeners, although the results are less outspoken for this vowel contrast. 95 and 91 of the 96 /hʏs/ stimuli (or 99.0% and 94.8%) were categorised accurately by Ghent and Antwerp listeners, respectively; for /hys/, Ghent and Antwerp listeners categorised 87 and 74 (90.6% and 77.1%) of the 96 stimuli accurately, respectively.

Figures 5.3, 5.4 and 5.5 above reveal why the Standard Dutch /i-ɪ/ contrast yields more striking results for Antwerp listeners than the Standard Dutch /y-ʏ/ contrast. The average spectral distance between the vowels of the target stimuli in the latter vowel pair is larger, and the durational differences are far more outspoken. While the shorter vowels of the two pairs (/ɪ/ and /ʏ/) are comparable in duration, /y/ is produced almost twice as long as /i/. That Antwerp listeners still make relatively more mistakes in the categorisation of /y/ and /ʏ/ indicates a discrepancy between Antwerp and Ghent listeners (in spite of the p-values in Table 5.7 indicating no significant differences).

To assess this discrepancy, we will examine which target vowels were chosen in the case of incorrect categorisation. With respect to Standard Dutch /y/, the incorrect mappings in the two groups of listeners are comparable, with the stimuli being categorised mostly as /ø:/ (5.2 % for Ghent listeners, 14.6% for Antwerp listeners), followed by a categorisation as /ʏ/ (3.1 % for Ghent listeners, 5.2% for Antwerp listeners). With respect to Standard Dutch /ʏ/, Ghent listeners predominantly categorised this vowel correctly, and no mapping onto /ø:/ took place. This is not the case for Antwerp listeners: although only limited, 5 (5.2%) of /ʏ/ stimuli were categorised as /ø:/.

While the results for Ghent listeners are as expected, the results for Antwerp listeners are not exactly along the lines of what was hypothesised: no significant categorisation of the Standard Dutch /hys/ as /ʏ/ is seen (or at least not to such a clear extent as was the case for the categorisation of /i/ as /ɪ/ by Antwerp listeners). Furthermore, no categorisation of the /hys/ stimuli as /ø:/ was predicted, but the results show that such categorisation is in fact even more frequent than that of /hys/ as /ʏ/.

When looking at this issue in more detail, some explanations arise which do not have to contradict the Antwerp listeners' results. The categorisation of both /y/ and /ʏ/ as /ø:/ could indicate that, in the phonetic categories constructed by Antwerp listeners, the

spectral features of this vowel pair are closer to each other than is the case for Ghent listeners, where only /y/ is categorised as /ø:/. If we consider the spectral vowel features mapped by Verhoeven & Van Bael 2002a: 157 (Figure 4.5 in the present study), we see that the Antwerp realisation of /ø:/ (transcribed as 'O' in the figure) lies close to the East-Flemish and Limburg (and, by inference, Standard Dutch) realisation of /ʏ/, which could explain that Antwerp (but not Ghent) listeners map /ʏ/ onto /ø:/. The accurate categorisation of /ʏ/ by Antwerp listeners could be attributed to the same explanation that applies to their relatively accurate /hɪs/ categorisation: Standard Dutch /ʏ/ could be considered a *new* vowel to listeners adjusted to an Antwerp production environment, facilitating its perception.

That more /hys/ stimuli are inaccurately categorised as /ø:/ than as /ʏ/, even though Standard Dutch /y/ lies spectrally farther from Antwerp /ø:/ than from Antwerp /ʏ/, could be explained by the fact that durational features play a significant role in the vowel perception. Since Standard Dutch /y/ was pronounced very long in comparison with Standard Dutch /ɪ/, more confusion arose with regard to another long vowel in the same vowel space, i.e. /ø:/. This finding would indicate that Antwerp listeners display a tendency towards attaching more importance to durational features than to spectral features (at least for the vowels under consideration). That more /y/ stimuli were mapped onto /ø:/ by Antwerp than by Ghent listeners, could also be due to this relatively higher importance attached to the durational feature by Antwerp listeners, which would lead these listeners to map vowels onto a different vowel with equal duration (even if there are spectral discrepancies). In Escudero's L2LP (2005) terms, the more developed durational cue constraints in Antwerp listeners would cause them to put comparatively less weight on spectral cue constraints in their vowel perception.

The above discussion revealed that there are indeed regional differences in the L1 perception of Standard Dutch vowels by Ghent and Antwerp listeners, be it more outspoken

for some vowels than for others. Sections 5.5.5 and 5.5.6, which conclude the discussion of the perception experiment, will assess whether these regional differences are maintained cross-linguistically, i.e. when mapping RP English stimuli onto L1 Dutch vowel categories.

5.5.5 Cross-language perception task: /i:/ and /ɪ/

The results for the cross-language perception task regarding the /i:-ɪ/ contrast correspond to the hypotheses formulated in section 5.3.3. What stands out is that the overall performance in the categorisation of this vowel contrast is more accurate[20] than the categorisation of the Standard Dutch /i-ɪ/ contrast. This is even the case for Ghent listeners: while their performances for Dutch and English /ɪ/ are comparable (94 or 97.9% and 95 or 99.0% correct categorisations, respectively), their performance for English /i:/ is more accurate (95 or 99.0% correct categorisations) than their performance for Standard Dutch /i/ (86 or 89.6% correct categorisations). The Antwerp listeners perform slightly better in categorising English /ɪ/ as /ɪ/ than Standard Dutch /ɪ/ (92 or 95.8% and 86 or 89.6% correct categorisations, respectively), and perform notably better in categorising English /i:/ as Dutch /i/ than its Standard Dutch counterpart /i/ (82 or 85.4% vs. 51 or 53.1% correct categorisations, respectively). This is also reflected in the p-values presented in Table 5.10, where the difference for /i:/ categorisations is significant, unlike the difference for /ɪ/ categorisations.

These results are not surprising when we consider Figures 5.3, 5.4 and 5.5. While spectrally the F2 difference between the members of the Dutch /i-ɪ/ and the English /i:-ɪ/ pair is comparable, the F1 difference is much more outspoken, facilitating correct perception. Concerning the durational difference within the Dutch and English contrasts, both Standard Dutch /i/ and RP English /i:/are longer than their respective /ɪ/ complements. Nevertheless, the durational contrast is more outspoken for the English

[20] When English /i:/ and /ɪ/ are seen as phonemically equivalent to Dutch /i/ and /ɪ/, respectively.

target stimuli than for the Dutch target stimuli: the average duration of the Standard Dutch /i/ stimuli is 1.46 times the average duration of the Standard Dutch /ɪ/ stimuli, while the average duration of the RP English /i:/ stimuli is 1.68 times the average duration of the RP English /ɪ/ stimuli. This greater durational contrast (together with the larger absolute duration of the /i:/ stimuli) thus helps Antwerp listeners (and apparently also Ghent listeners) to categorise English /i:/ as Dutch /i/ more often than Standard Dutch /i/. RP English thus possesses vowel features that, when passing for Standard Dutch, facilitate a correct performance in categorisation. The explanation that the listeners had successfully constructed phonetic categories to accurately perceive these L2 vowels is not likely, precisely because the participants were led to believe that the L2 target stimuli were L1 vowels.

Although the overall performance of both groups is good, regional differences which were also present in the Standard Dutch perception task can still be discerned. While only 1 of the 96 RP English /i:/ stimuli was categorised as /ɪ/ by Ghent listeners, Antwerp listeners still categorised 13 (or 13.5%) of English /i:/ stimuli as /ɪ/. Although only to a limited extent, Antwerp listeners again categorised 3 of the /ɪ/ stimuli as /ɛ/, whereas Ghent listeners never chose this mapping option. These two findings, which correspond to patterns also seen in the Dutch perception tasks, indicate that there are subtle regional differences in the perception of L2 vowels (at least when these are thought to be L1 vowels).

The fact that Antwerp listeners perform notably more accurate on the English task than on the Standard Dutch task leaves us with a surprising implication. If listeners are confronted with L2 vowels with features that correspond better to their own dialectal background than L1 standard language vowels do, the perceptual performance on an L2 vowel contrast can be better than the performance on a corresponding contrast in the standard variety of their own native language.

5.5.6 Cross-language perception task: /ʊ/ and /ʌ/

Concerning RP English /ʊ/, the hypothesis that this vowel would mostly be classified as Dutch /u/ is confirmed. Ghent listeners use /u/ the most to map English /ʊ/ onto: 94 or 97.9% of the 96 /ʊ/ stimuli are categorised as such. Antwerp listeners, in comparison, do not choose /u/ as often as Ghent listeners: 81 or 84.4% of the stimuli are categorised as /u/ here. The difference in categorisation of /ʊ/ between the two dialect groups is significant (cf. Table 5.10).

An analysis of the rest of the other vowels /ʊ/ was mapped onto indicates some regional perceptual differences. The Antwerp listeners categorise 14 or 14.6% of the 96 /ʊ/ stimuli as /ʏ/, as opposed to the Ghent listeners who categorise none of the /ʊ/ stimuli as such. This confirms the hypothesis, which predicted that because of the spectral and durational features of /ʊ/ being more similar to Antwerp /ʏ/ than to Ghent /y/, equivalence classification/perceptual assimilation could take place for Antwerp listeners, but not for Ghent listeners.

Furthermore, Ghent listeners categorise /ʊ/ as /y/ (2 of the 96 stimuli), while Antwerp listeners do not. Since only 2 such mappings occur, caution is needed when deriving that Ghent /y/ is relatively shorter perceptionally than Antwerp /y/, as Ghent listeners use their /y/ to substitute for /ʊ/, whereas Antwerp listeners do not do so.

With regard to the perception of RP English /ʌ/, the hypothesis that this vowel is categorised predominantly as /a:/ or /ɑ/ (and to a lesser extent as /ʏ/) in both listener groups is confirmed. Although the Ghent and Antwerp participant groups chose the same three vowel categories to map this vowel onto, the regional difference for this stimulus is significant ($p = 0.032 < 0.05$, cf. Appendix D), indicating a different distribution into these categories. In Ghent listeners, a roughly equal distribution can be seen between /a:/ and /ɑ/ (42 or 43.8% and 38 or 39.6%, respectively of the 96 stimuli). Antwerp listeners rather

display a preference towards a mapping onto /ɑ/ (55 or 57.3% of the 96 stimuli, vs. 33 or 34.4% categorisations as /a:/). This pattern was predicted by the hypothesis: the spectral coinciding of /a:/ and /ɑ/ in Antwerp (Brabantine) Dutch (as shown in Figure 4.5) could explain why Antwerp listeners are more inclined to choose the shorter alternative of the two vowels upon perceiving a relatively short vowel (cf. Figure 5.4). A greater attention to the feature of vowel duration in Antwerp listeners could explain this observation as well. Figure 4.5 furthermore reveals that there are spectral differences between East-Flemish /a:/ and /ɑ/. Specifically, English /ʌ/ coincides with East-Flemish /a:/ on a spectral level, and with East-Flemish /ɑ/ on a durational level, explaining the equal distribution between the two alternatives. Although the participants were not aware that the stimuli were of an L2, this situation could be seen as an instance of Escudero's proposed SUBSET scenario of L2 perception: indeed, "the L1 perception grammar outputs more categories than required in the L2" (Escudero, 2005: 123).

Lastly, the hypothesis that the Ghent listeners will categorise English /ʌ/ as /ʏ/ to a greater extent than Antwerp listeners is also confirmed (16 or 16.7% vs. 8 or 8.3% of the 96 stimuli, respectively). Antwerp and Ghent /ʏ/ both are the short counterpart of the /y-ʏ/ contrast, which suggests that the smaller number of mappings onto /ʏ/ by Antwerp listeners finds its origins on a spectral level: since Antwerp /ʏ/ is more closed than Ghent /ʏ/, it appears to be a less likely candidate for perceptual assimilation.

In conclusion, the analysis of this perception experiment has shown that there are indeed regional differences in the perception of L1 Standard Dutch vowels by Ghent and Antwerp listeners, and that these differences appear to play a role in the perception of L2 vowels as well. Whether these regional differences are also apparent in the L1 and L2 production of the two groups of participants is studied in the following section, which presents the production experiment.

6 Production experiment

6.1 Aim

The results from the perception experiment have indicated influence of dialectal features on the perception of a number of Standard Dutch and RP English vowels. The aim of the present experiment is twofold. By conducting a Dutch picture-naming and sentence-reading task, a first part examines whether the different spectral and durational vowel features of the L1 /i-ɪ/ and /y-ʏ/ contrasts in Ghent (East-Flemish) and Antwerp (Brabantine) Dutch (cf. section 4) can also be detected in productions of the present study's participants.

Subsequently, a similar English picture-naming and sentence-reading task examines whether this presumed L1 regional variation exerts an influence on the production of English /iː ɪ ʊ ʌ/. The production data of the two participant groups are compared to each other (statistical analyses are performed to assess the significance of the results) as well as to Dutch production data from Verhoeven & Van Bael (2002b) and RP English data from speakers of the same age group, taken from Hawkins & Midgley (2005). The results will reveal whether there are any significant discrepancies in the L2 productions of the two speaker groups that could be accounted for by L1 regional variation.

6.2 Methodology

6.2.1 Participants

As the participants are the same for the perception and the production experiments, we refer to section 5.2.1 for the relevant information and discussion.

6.2.2 Design of the task and stimuli

The production experiment consists of three parts: a Dutch task, an oral questionnaire (conducted in English), and an English task. The purpose of inserting the questionnaire (which was necessary to help account for possible intra-speaker discrepancies) in the middle was that, by conducting it in English, the informants were prepared for the second task: their 'English mode', as it were, was activated.

As mentioned in section 4, Verhoeven & Van Bael (2002a&b) essentially studied regional variants of Standard Dutch production; for this reason, their task setup had to be formal, which is why non-words with an identical consonantal context in a fixed carrier sentence were used. While the tasks in the present study also had to be of a certain formality in order to make comparison possible, more spontaneous, regional language production was aimed at (especially for the Dutch task). It was thus decided that it would be disadvantageous to work with non-words in identical carrier sentences and in identical consonantal environments. For this reason, the Dutch and the English tasks both consisted of a picture-naming as well as a sentence-reading part, which made use of meaningful words in varying carrier phrases. Despite the inconvenience of lexical variation concerning vowel production, existing words were chosen over non-words. That the production tasks were nevertheless considerably formal contributes to the strength and relevance of the potential differences.

The purpose of repeating the target stimuli in a carrier sentence (inserted after each picture) was firstly to rule out potential influence of rising intonation (related to the

guessing of the words in the picture-naming task), and secondly to create a less formal context (in comparison with the picture-naming task). First, a slide was presented which showed only the picture. After the participant had guessed the target stimulus, the slide was repeated. This second slide contained the same picture, but now with the carrier sentence displayed underneath the picture. Participants had to insert the target stimulus in the blank, indicated by three dots (as shown in Figure 6.1). Furthermore, the target stimuli all occurred at the end of declarative sentences, so that it was ascertained that they were all produced with a similar intonation pattern.[21]

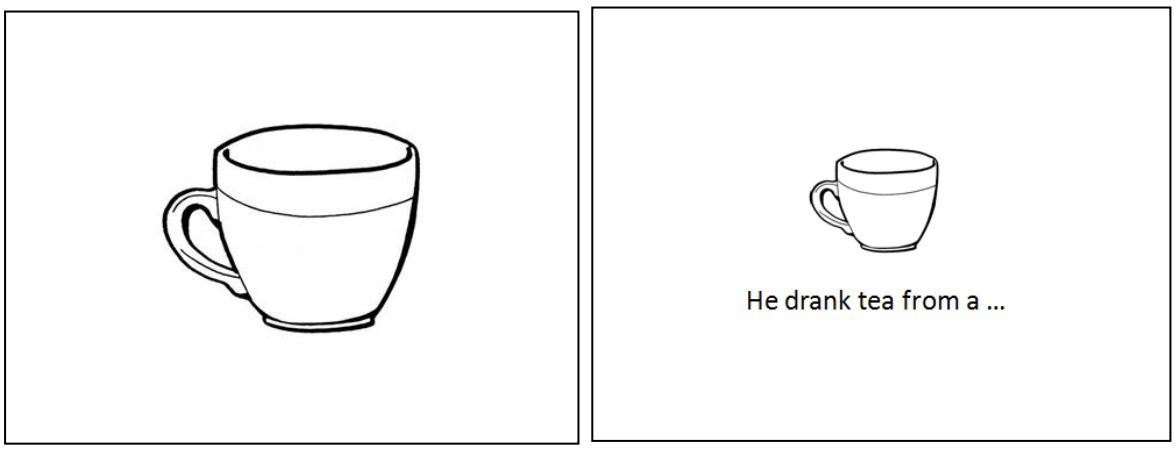


Figure 6.1. An example pair of slides taken from the English task.

The actual Dutch and English tasks are both preceded by an introductory two-word test trial, to guarantee familiarity with the exact procedure. Four target stimuli per vowel are used, providing a total of 16 stimuli per informant per task. Voiceless consonantal environments were chosen whenever possible, and velarised /l/ ([ɬ]) at the end of words was avoided because of its influence on the quality of the preceding vowel. However, the need for concrete stimuli that could easily be identified and named limited the available choices, and compromises had to be made in several cases (cf. section 6.4.1 below).

[21] The only exception was the English 'spit' stimulus ('there was spit on the floor.'), since the marked sentence structures that would arise from a sentence-final position of this stimulus were judged to be disadvantageous to a spontaneous production.

Nine distracters were added to each task, providing a total of twenty-five words per task. The order of the pictures was then randomised. Tables 6.1 and 6.2 present an overview of the target stimuli and distracters. The visual stimuli and carrier sentences for the Dutch and the English production task are presented in part 1 and part 2 of Appendix E, respectively.

Table 6.1. The target stimuli and distracters included in the Dutch picture-naming and sentence-reading task.

Dutch			
Target stimuli			
/i/	/ɪ/	/y/	/ʏ/
kies /k i s/ 'molar'	schip /s x ɪ p/ 'ship'	muur /m y r/ 'wall'	put /p ʏ t/ 'pit'
brief /b r i f/'letter'	sik /s ɪ k/ 'goatee'	stuur /s t y r/ 'steering wheel'	schub /s x ʏ p/ 'scale'
dief /d i f/ 'thief'	pit /p ɪ t/'kernel'	duur /d y r/'expensive'	kus /k ʏ s/ 'kiss'
ziek /z i k/ 'ill'	kist /k ɪ s t/ 'chest'	vuur /v y r/'fire'	brug /b r ʏ x/'bridge'
Distracters			
beer /b eː r/, 'bear'	kaas /k aː s/, 'cheese'	glas /ɣ l ɑ s/, 'glass'	
bal /b ɑ l/, 'ball'	stoel /s t u l/, 'chair'	muis /m œ y s/, 'mouse'	
aap /aː p/ 'monkey'	geel /ɣ eː l/, 'yellow'	huis /h œ y s/ 'house'	

Table 6.2. The target stimuli and distracters included in the English picture-naming and sentence-reading task. Since British English is used as the target norm (cf. the preliminary note on American and British English in section 4), British RP English transcriptions are used in this table).

English			
Target stimuli			
/iː/	/ɪ/	/ʊ/	/ʌ/
key /k iː/	spit /s p ɪ t/	book /b ʊ k/	hug /h ʌ g/
sheep /ʃ iː p/	fist /f ɪ s t/	foot /f ʊ t/	duck /d ʌ k/
peace /p iː s/	six /s ɪ k s/	hook /h ʊ k/	cup /k ʌ p/
leaf /l iː f/	pig /p ɪ g/	cook /k ʊ k/	bus /b ʌ s/
Distracters			
dog /d ɒ g/	cat /k æ t/	bat /b æ t/	
bed /b ɛ d/	red /r ɛ d/	grass /g r ɑː s/	
door /d ɔː/	rose /r ə ʊ z/	shoe /ʃ uː/	

6.2.3 Procedure

The production experiment was conducted in the same session as the perception experiment, in a quiet, well-isolated room. The productions were recorded with a Marantz PMD620 digital recorder and a Sony ECM-MS907 one-point stereo microphone.

The informants were briefly explained how the task was going to be conducted. Due to the two-word test trial at the beginning of each task, the relative clearness of the pictures and the relatively simple concepts they represented, difficulties in naming the pictures were avoided as much as possible.

6.2.4 Coding and analysis

The 12 Ghent and 12 Antwerp participants each produced 4 productions of each of the 4 stimuli in each of the two production tasks, for a total of 384 Dutch and 384 English productions. The task results were analysed using Praat (version 5.3.45, Boersma & Weenink, 2013). The vowel duration of the stimuli was determined through auditory analysis of the sound files and visual analysis of the accompanying acoustic waveforms, and the vowel boundaries of each target vowel were marked. Subsequently, the exact middle of the vowel was determined automatically by Praat (this point was chosen to ensure a minimal influence of consonantal context). At this midpoint, a formant measurement was conducted, by which the first and second formant frequency (F1 and F2, respectively) were determined. No script was used for this: manually analysing each vowel ensured a better accuracy and made it possible to reconsider values that could not be correct. Figure 6.2 below illustrates how formant analysis with the use of scripts could yield distorted results. The Praat interface of the /ki:/ production by Antwerp participant 10 shows several formant values (visualised by red dots) that deviate from the otherwise consistent line of red dots indicating the actual F2 value of the vowel in question. In Figure 6.2, the F2 value at the deviating point is approximately 1077 Hz, while the (correct) value at the midpoint is

2692 Hz. In cases where the situation is reversed, and a distorted result would be observed at the midpoint of the vowel as determined by Praat, it goes without saying that this would heavily influence the results of the tasks. Therefore, manual measurement was preferred over the use of measurement scripts, so that the position at which formant measurements took place could be slightly adjusted, whenever this was deemed necessary.

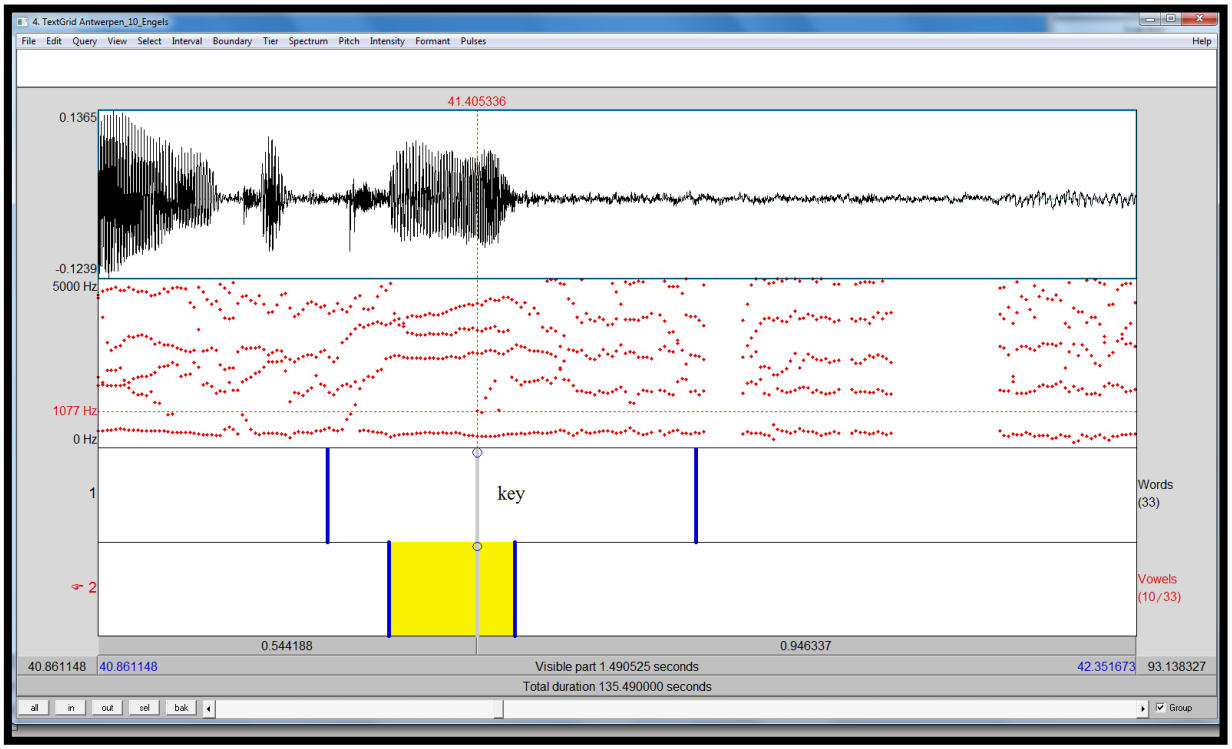


Figure 6.2. Example of the Praat interface (for the /ki:/ target stimulus produced by Antwerp participant 10). The red dotted axis indicates where automatic formant measurement would be problematic if it occurred at the Praat-determined midpoint.

For reasons already mentioned in section 5.2.4, only the first two formants were measured; moreover, the normalisation method used in the present study only requires information on the first two formants. As this vowel normalisation proved to be a beneficial approach in Debaene (2012) with regard to the visual representation of the results, the Lobanov (1971) normalisation method is also applied in the present study (cf. section 6.4.2.2 below).

To assess whether there are significant between-dialect differences, the task results were entered in the SPSS (IBM Corp., 2012, version 21.0) tool for statistical analysis. To make sure sex-based differences were not biasing the results, male and female participant results

were separated for the statistical tests.[22] That the male sample then becomes very small (2 Antwerp and 4 Ghent male participants) will be kept in mind in the presentation and discussion of the results. However, by conducting the statistical analysis in this fashion (and by controlling variables like education), we can assume that the only considerable variable is the city of origin, precisely the factor which is assessed in the present study.

An independent, two-tailed t-test for equality of means then examines whether the average F1, F2 and duration results for the Antwerp participants are significantly different from those of the Ghent participants for a particular sound. The (statistical) hypotheses of the t-tests are as follows:

- Null hypothesis (H_0): there is no significant difference between the averages.
- Alternative hypothesis (H_A): there is a significant difference between the averages.

In this specific case, the null hypothesis indicates that there is no significant difference between Antwerp and Ghent productions. This hypothesis is rejected if the p-value is smaller than (the generally accepted and predetermined) 0.05 or 5%, corresponding to a result that is unlikely to be based on arbitrary chance.

For the t-test, the p-value depends on the outcome of the Levene's test, which calculates whether or not the variances (i.e. the standard deviation squared) of the Ghent and Antwerp vowel data are equal. Again, there are two hypotheses:

- Null hypothesis (H_0): there are equal variances.
- Alternative hypothesis (H_A): there are significantly differing variances.

[22] The size of the vocal tract (which differs between males and females) can influence the formant frequencies; a generally larger vocal tract in men generates lower formant frequencies, while a comparatively smaller vocal tract in women encompasses higher formant frequencies (Verhoeven & Van Bael, 2002 a&b) - cf. also section 6.4.2.2, in which vowel normalisation is treated.

For this test, too, a p-value of 0.05 or less is needed to reject the null hypothesis. If this is the case, a correction of the p-value of the initial t-test is needed to avoid a distorted result that follows from rejecting the null hypothesis, i.e. that the variances are not equal.

To conclude the present discussion, two examples illustrate how the above theory is put into practice. Figure 6.3 shows the SPSS output for the durational feature for female Antwerp (n= 10) and Ghent (n= 8) participants with regard to the Dutch 'sik' (/s ɪ k/) stimulus or 'SikV.L' ('V' refers to 'vrouwelijk', Dutch for 'female'). Levene's test indicates significantly differing variances between the two participant groups; the (corrected) p-value (or *'sig. (2-tailed)'*) from the second row is thus selected in this particular case. The p-value here is then 0.553; since this value is higher than 0.05, the null hypothesis cannot be rejected, and we can conclude that, on the basis of the tested sample, there is no significant difference in average duration for the female productions of this stimulus. This corresponds to the predictions in section 5.3.1: it was expected that both Antwerp and Ghent speakers would produce /ɪ/ as a short vowel.

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
SikV.L	1,00	10	67,7200	7,81733	2,47206
	2,00	8	70,8875	12,84784	4,54240

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
SikV.L	Equal variances assumed	4,935	,041	-,647	16	,527	-3,16750	4,89725	-13,54921	7,21421
	Equal variances not assumed			-,612	11,009	,553	-3,16750	5,17150	-14,54872	8,21372

Figure 6.3. SPSS output for the durational feature of the female 'sik' (/s ɪ k/) productions. The Levene's test and its p-value ('Sig.') are indicated in red, the t-test and its uncorrected (1st row) and corrected (2nd row) p-values are indicated in green.

Figure 6.4 shows the SPSS output for the feature of duration for male Antwerp (n= 2) and Ghent (n= 4) participants with regard to the Dutch 'kies' (/k i s/) stimulus or 'KiesM.L'.

The p-value for the Levene's test indicates equal variances in both groups (p= 0.861 > 0.05); the uncorrected p-value of the t-test is selected in this case, producing a highly significant result (p= 0.006 < 0.05). Although the sample is relatively small, this result also corresponds to the predictions: regarding /i/ duration, a significant difference between Antwerp and Ghent speakers was expected. More detailed statistical results are presented in sections 6.4 and 6.5 below, which provide a presentation of the results of the production tasks, and in Appendices I and J, which present the SPSS outputs per target vowel, and the p-values per target stimulus (for male and female F1, F2 and duration values), respectively.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
KiesM.L	1,00	2	129,8000	13,57645	9,60000
	2,00	4	74,7250	11,59062	5,79531

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
KiesM.L	Equal variances assumed	,035	,861	5,248	4	,006	55,07500	10,49417	25,93851	84,21149
	Equal variances not assumed			4,911	1,783	,049	55,07500	11,21364	,72979	109,42021

Figure 6.4. SPSS output for the durational feature of the male 'kies' (/k i s)/productions. The Levene's test and its p-value ('Sig.') are indicated in red, the t-test and its uncorrected (1st row) and corrected (2nd row) p-values are indicated in green.

It goes without saying that the present study concerns a statistical sample. Although in total, a considerably large number of vowels was measured (384 Dutch and 384 English productions), the analysed number of participants in each dialect group was still relatively small (12 Ghent and 12 Antwerp participants). Irregularities in the data are always possible, so caution is required when extrapolating the conclusions from the present study to the whole Antwerp and Ghent population, respectively.

6.3 Hypotheses

6.3.1 Dutch Production task

Concerning the Dutch production task, the hypothesis is that the results will correspond to the Ghent and Antwerp vowel systems as described in section 4. Antwerp speakers are thus predicted to have a realisation of the /i-ɪ/ and /y-ʏ/ contrast in which the vowels are relatively closer on a spectral level, and more distinct on a durational level. The Ghent speakers, on the other hand, are predicted to realise the vowels of these contrasts as more distinct on a spectral level, and more equal on a durational level.

6.3.2 English Production Task

At least some influence of Antwerp Brabantine Dutch is to be expected when speakers pronounce the English vowels /iː ɪ ʊ ʌ/, especially since the spectral characteristics of Antwerp Brabantine Dutch (keeping its expansive nature in mind, cf. section 4.1.3) are relatively unchallenged by Standard Dutch. These characteristics are commonly transferred when speaking Standard Dutch, as was pointed out in section 4.1.3, suggesting that stable phonetic L1 categories have been constructed for the Antwerp Brabantine Dutch variants of the vowels. The results from the perception experiment in the present study supported this observation. The Ghent (East-Flemish) Dutch speakers are hypothesised to have both an L1 and L2 production that is less deviating from that of Standard Dutch and RP English than the Antwerp speaker productions, as we can derive from the discussion in section 4. This was also proven on a perceptual level in the previous experiment, where Ghent participants had greater accuracy rates for the categorisation of all Standard Dutch and RP English vowels.

The discrepancy between Antwerp (Brabantine) and Ghent (East-Flemish) Dutch would, according to the theoretical models forwarded in section 2, suggest that differences in L2 production are likely, since L2 perception (and the subsequent production) in these

models depends on the L1 (production) background. Concerning English /i:-ɪ/, it is hypothesised that production of the spectral differences between the members of this pair will prove more problematic for the Antwerp speakers, while Ghent listeners may have more difficulty with the durational distinction (under the provision, of course, that optimal L2 perception and production have not yet been obtained). If we are to follow Escudero's L2LP (2005), English /i:/ and /ɪ/ can be seen as *similar* to Antwerp /i-ɪ/ and Ghent /i-ɪ/: the mapping of the sounds is non-optimal. Escudero's L2LP (2005) predicts only one learning task, namely that the category boundaries be shifted in order to accurately perceive the L2 contrast. Ghent speakers will have to adjust their categories to include a durational distinction, while Antwerp speakers will need to do the same to distinguish these two English vowels on a spectral level. As Escudero asserts that "problems producing L2 sounds originate in large measure from difficulties in perceiving such sounds in a native-like fashion" (2005: 325), it is thus hypothesised that, due to this regional difference in L2 perception, Antwerp speakers will have more difficulties with the spectral distinction of English /i:/ and /ɪ/, while Ghent speakers will have more difficulties with the durational distinction of English /i:/ and /ɪ/. This SIMILAR L2LP (Escudero, 2005) scenario is predicted to be relatively less difficult than if it were to be seen as a NEW contrast. If we assume a (simplified) total merging of /ɪ/ and /i/ in Antwerp Brabantine Dutch, English /ɪ/ could be seen as NEW, as it is rather different spectrally from this L1 contrast. In this case, more difficulties are predicted in Escudero's L2LP (2005) than for a SIMILAR vowel, since more learning tasks are involved.

In Flege's SLM (1995) a different prediction emerges with regard to the latter scenario: an L2 vowel will be easier to learn if it is *new* (no interference occurs with L1 phonetic categories). For Antwerp speakers, *new* /ɪ/ should therefore be rather easy to produce, as a new phonetic category needs to be created for this sound. If this English /ɪ/ is to be seen as *similar* to Antwerp /ɪ/, the SLM (Flege, 1995) predicts that the ultimate

attainment of the vowel is rather difficult due to potential equivalence classification. Through this process, English /ɪ/ could be produced more closed and fronted, and thus /i/-like (similar to its Antwerp area counterpart), which moreover could give rise to confusion between English /ɪ/ and /iː/.

With regard to RP English /ʊ/, equivalence classification to Dutch short /u/ is expected for both groups, as this is the only vowel in this vowel region to which English /ʊ/ is expected to be perceptually assimilated. However, since the vowel space where /ʊ/ is generally fronted to in English (of younger speakers, cf. Hawkins & Midgley, 2005: 188) is already occupied by a long and a short vowel in Antwerp (Brabantine) Dutch (/y/ and /ʏ/), the prediction is that the Antwerp productions of /ʊ/ will be less fronted than the Ghent productions. Durationally, Antwerp productions could prove to be longer than Ghent productions if /ʊ/ is perceptually assimilated to Dutch /u/: Antwerp /u/ is produced almost twice as long as its East-Flemish counterpart, according to Figure 4.5 in section 4.1.3.

Concerning RP English /ʌ/, it is hypothesised that the Ghent productions will lie in the vowel space between Dutch /aː/ and /ʏ/ due to perceptual assimilation to one of these two vowels. If /ʌ/ is seen as a *new* vowel, accurate production could take place as well. As Antwerp /ʏ/ is spectrally almost similar to Antwerp /y/, and subsequently more closed than Ghent /ʏ/, the hypothesis here is that equivalence classification of English /ʌ/ to Dutch /ʏ/ is less likely to occur in Antwerp (Brabantine) Dutch. With respect to durational features, this vowel is expected to be produced short.

6.4 Results

6.4.1 Preliminary note: lexical variation

In Debaene (2012), two of the target stimuli had to be removed from the overall results. The average durational features of Dutch target stimulus 'fiets' (/f i t s/, 'bicycle') was even shorter than the average for /ɪ/[23], which is why this stimulus was avoided in the present study. Furthermore, the deviating spectral values of the English target stimulus 'shell' (caused by the lowering and centralising effect of [ɬ], velarising or 'dark' /l/) made the values of this vowel unsuitable for cross-linguistic comparison in Debaene (2012), which is why [ɬ] was avoided in word-final positions in the present study. However, the use of meaningful words in the present study still restricted the choice of stimuli. In some cases, compromises had to be made, the most important of which are presented here.

It is generally acknowledged that vowels are produced longer before voiced consonants than before voiceless consonants (Chen, 1970; Kluender, Diehl & Wright, 1988). Because of this, it was necessary to devise target stimuli that all ended either in a voiceless or in a voiced consonant. To facilitate the visual discernment of the (end) boundaries of the vowels, target stimuli ending in voiceless consonants were used whenever possible. Since the choice of monosyllabic stimuli was limited, two exceptions had to be made for the English task, namely for 'pig' /p ɪ g/ and 'hug' /h ʌ g/. When we look at the average vowel duration of these two target stimuli in both participant groups, we see that 'pig' and 'hug' are indeed produced longer than the other stimuli for English /ɪ/ and /ʌ/, respectively (except Antwerp 'hug' - cf. Tables 6.3 and 6.4). However, the durational differences were not as outspoken as between 'fiets' and the other Dutch /i/ stimuli from Debaene (2012). Therefore, it was decided not to exclude these stimuli from the final results.

[23] We do not at present have an explanation for the deviating vowel length in the word 'fiets', cf. Debaene (2012).

Tables 6.3 and 6.4 present the average duration, standard deviation and median of the /ɪ/ and /ʌ/ stimuli for Ghent and Antwerp participants.

Table 6.3. Average duration, standard deviation and median of the 4 /ɪ/ and /ʌ/ stimuli for Ghent speakers.

Ghent			
Stimulus	Average length (ms)	Standard deviation (ms)	Median (ms)
spit /s p ɪ t/	73.8	10.2	72.3
fist /f ɪ s t/	82.1	13.7	82.2
six /s ɪ k s/	80.1	14.3	78.3
pig /p ɪ g/	113.5	20.3	107.3
Stimulus	Average length (ms)	Standard deviation (ms)	Median (ms)
duck /d ʌ k/	105.1	17.2	100.4
cup /k ʌ p/	94.8	13.7	91.3
bus /b ʌ s/	105.3	17.5	108.6
hug /h ʌ g/	118	25.9	113.1

Table 6.4. Average duration, standard deviation and median of the 4 /ɪ/ and /ʌ/ stimuli for Antwerp speakers.

Antwerp			
Stimulus	Average length (ms)	Standard deviation (ms)	Median (ms)
spit /s p ɪ t/	68.1	10.0	70.3
fist /f ɪ s t/	75.1	9.0	76.4
six /s ɪ k s/	78.0	9.2	78.5
pig /p ɪ g/	94.1	10.2	94.7
Stimulus	Average length (ms)	Standard deviation (ms)	Median (ms)
duck /d ʌ k/	115.0	28.0	102.8
cup /k ʌ p/	89.7	13.4	89.0
bus /b ʌ s/	114.0	19.0	109.6
hug /h ʌ g/	110.7	11.7	108.8

Concerning the Dutch task, the /y/ stimuli did not correspond to the requirement of a voiceless final consonant. All stimuli ended in /r/, since virtually no Dutch monosyllabic words with /y/ exist that do not have /r/ as their final consonant (with only a few exceptions, such as some personal names like 'Guus', the water bird called 'fuut', and words ending in the /yu/ diphthong, like 'ruw' 'rough'). In their description of the Dutch close free vowels /i/, /y/ and /u/, Collins & Mees point out that "[b]efore /r/, all these vowels are

lengthened in all varieties of Dutch" (2003: 132). This observation is kept in mind in section 6.4.2.1, which presents the durational results.

It was decided that, in contrast with Debaene (2012), none of the target stimuli would be withheld from the results for two reasons. First, in the present study greater attention was paid to the selection of favourable consonantal environments. Second, and most importantly, Debaene (2012) could not include target stimuli with values that were deviating too much, since only one group (Antwerp speakers) was examined. The present study, on the other hand, is comparative; an identical procedure was followed for both speaker groups, and overall consonantal variation between their dialects is negligible[24]. In other words, it was ensured that none of the target stimuli displayed acoustic features that were so deviating as to justify excluding them from the results, and even if consonantal influence was involved, this applied to both speaker groups, so that between-group comparison was not jeopardised.

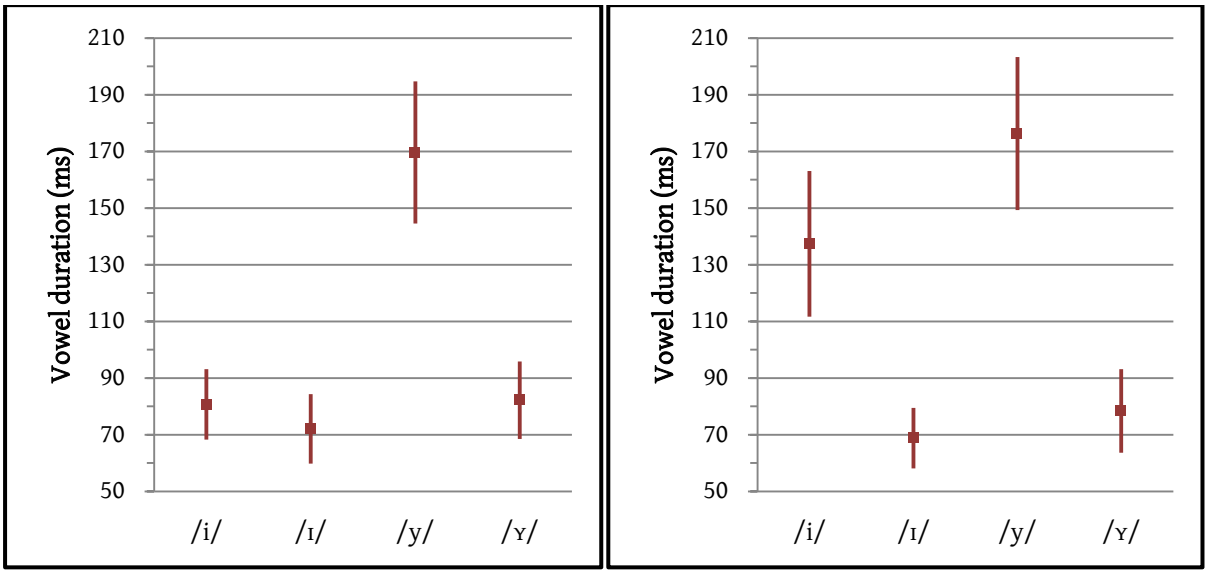
6.4.2 Overall results

The individual values of the two production tasks are presented (per participant group) in Appendix F. Appendix G presents the average values per target stimulus, grouped per production task and per participant group. Appendix H presents an overview of the production data per target vowel, grouped per production task and per participant group.

6.4.2.1 Durational features

Figures 6.5 and 6.6 show the average vowel durations of the target stimuli (grouped per target vowel) produced in the Dutch production task by Ghent and Antwerp speakers, respectively.

[24] One exception is /r/ (cf. section 2.2.3.3), of which the majority of Ghent speakers produced the uvular variant, while all Antwerp speakers produced the alveolar variant. If we keep in mind Collins & Mees' remark on /y/ preceding /r/ (2003: 132, as cited in the previous paragraph), it can be assumed that this allophonic variation is of no major influence on the results of the present study.



Figures 6.5 and 6.6. Average vowel duration (in ms) of the Dutch productions of the target stimuli by the Ghent (left) and Antwerp (right) participant group per target vowel. The vertical lines indicate the standard deviation range.

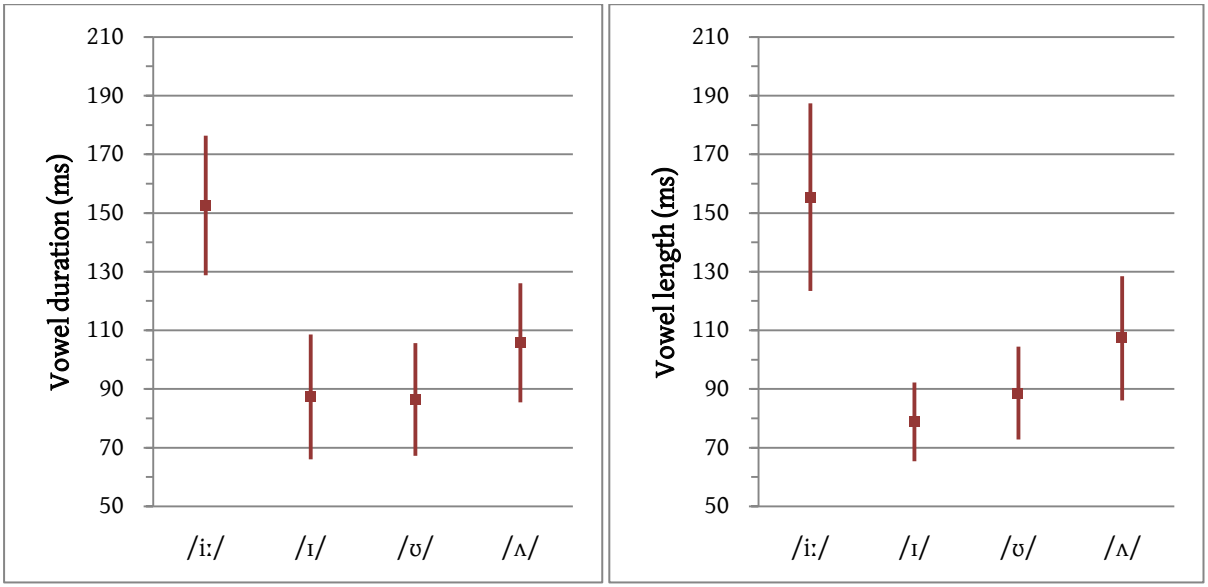
With regard to the Dutch productions of Ghent speakers, the results for the durational features largely correspond to the description in section 4.1.2. Indeed, the hypothesis that there is only a very subtle durational distinction in the Ghent /i-ɪ/ contrast is confirmed, with durational features that are very similar to each other (and to /ʏ/, which is also regarded as a short vowel). It is not surprising that the durational values of /y/, which is generally treated as a short vowel (as was pointed out in section 4), are rather high here, since /y/ precedes /r/ in the stimuli (cf. section 6.4.1).

Regarding the Dutch production of Antwerp speakers, the hypothesis for the /i-ɪ/ contrast is confirmed: whereas Ghent /i/ and /ɪ/ lie very close to each other durationally (with considerable overlap of standard deviations), Antwerp /i/ is, on average, twice as long as Antwerp /ɪ/, and there is no overlap between the respective standard deviations. Indeed, the average duration of Antwerp /i/ differs very significantly from that of Ghent speakers (both male and female $p < 0.001$ for duration).

The values of Antwerp /y/ are somewhat longer in comparison with those of Ghent /y/, and contrast somewhat more with the duration of /ʏ/ (as was expected). However, the

cross-dialectal difference is not as apparent as for the /i-ɪ/ contrast (an observation also seen in the perception experiment), and the durational mean values and standard deviations of this contrast are rather similar across the two dialects (/y/ duration: male p= 0.184 > 0.05 and female p= 0.382 > 0.05; /ʏ/ duration: male p= 0.066 > 0.05 and female p= 0.297 > 0.05; cf. Appendix I).

Figures 6.7 and 6.8 present the average vowel durations of the target stimuli (grouped per target vowel) produced in the English production task by Ghent and Antwerp speakers, respectively.



Figures 6.7 and 6.8. Average vowel duration (in ms) of the English productions of the target stimuli by the Ghent (left) and Antwerp (right) participant group. The vertical lines indicate the standard deviation.

For the English task, we see that Ghent speakers make a durational distinction between /i:/ and /ɪ/. The vowels /ʊ/ and /ʌ/ are relatively short, which was expected. As expected, Antwerp speakers make a slightly greater durational contrast between English /i:/ and /ɪ/ than Ghent speakers. When comparing averages, Ghent English /i:/ is produced 1.75 times longer than Ghent English /ɪ/ in the present study's sample, while Antwerp English /i:/ is produced 1.97 times longer than Antwerp English /ɪ/ in the present study's sample. Moreover, the difference between (mean value of English /i:/ minus 1 standard deviation) and (mean value of English /ɪ/ plus 1 standard deviation) is 20.2 ms for Ghent

speakers, and 31.2 ms for Antwerp speakers, implying a greater durational contrast in the latter group. Statistical analysis shows a significant difference for duration between the female English /ɪ/ productions ($p = 0.026 < 0.05$, cf. Appendix I), but no significant results are obtained for /iː/ (moreover, the t-test does not take into account between-vowel differences, i.e. differences in how vowels are contrasted within the two speaker groups).

With regard to English /ʊ/ and /ʌ/, the Antwerp results are parallel with the Ghent results on the 5%-significance level (for duration of /ʊ/, male $p = 0.417 > 0.05$ and female $p = 0.911 > 0.05$; for duration of /ʌ/, male $p = 0.287 > 0.05$ and female $p = 0.614 > 0.05$; cf. Appendix I). With regard to /ʌ/, this does not contradict the hypothesis, but the predicted potential influence of a much longer Dutch /u/ vowel in Antwerp (Brabantine) Dutch is not observable nor is it statistically significant.

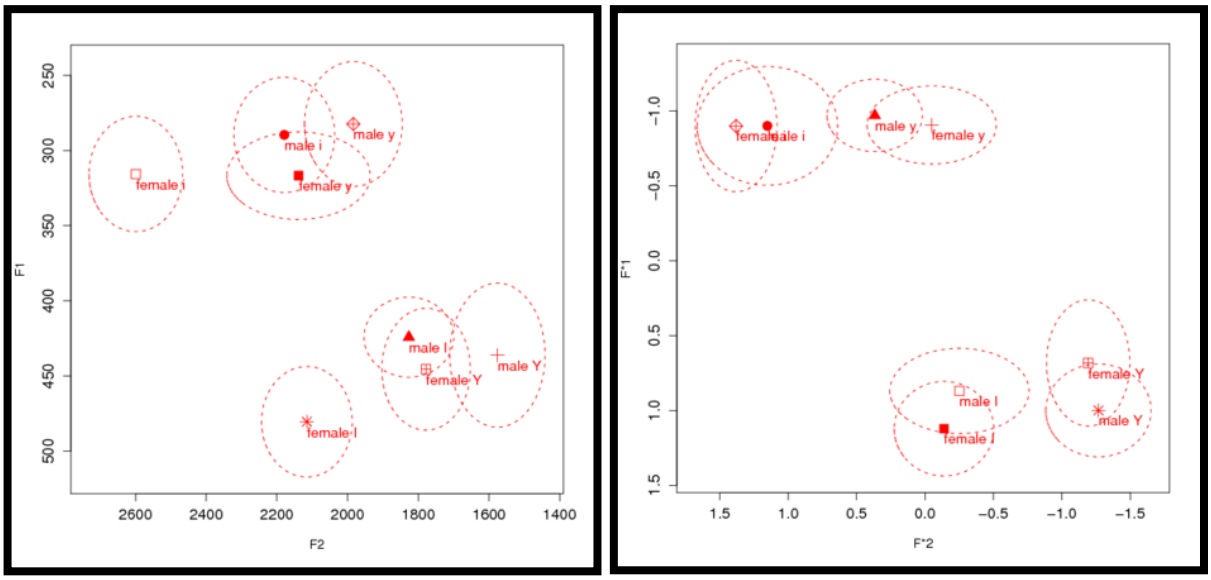
6.4.2.2 Spectral features

Within each speaker group, there were more female than male participants. Moreover, the ratio of male to female participants was different in the two groups (4 male participants out of 12 for the Ghent group, 2 male participants out of 12 for the Antwerp group). These two factors make it highly unfavourable to use the non-normalised production data in a visual representation of the data and subsequent analyses, since sex-related physiological differences in the size of vocal tracts can largely influence the formant frequencies (Verhoeven & Van Bael, 2002a: 149 & 160). This is also why male and female results were separated in the statistical analysis. Vowel normalisation can solve these problems, and it can provide a more satisfying visual representation of the data. The Lobanov normalisation method as devised by Lobanov (1971) was opted for in the present study, since Verhoeven & Van Bael (2002a&b) have proven that a similar method is effective in eliminating physiological sex-related differences, while preserving regional differences. Moreover, a comparison of vowel normalisation methods in Adank *et al.* examined "how effectively [the normalisation procedures] (a) preserve phonemic information, (b) preserve information

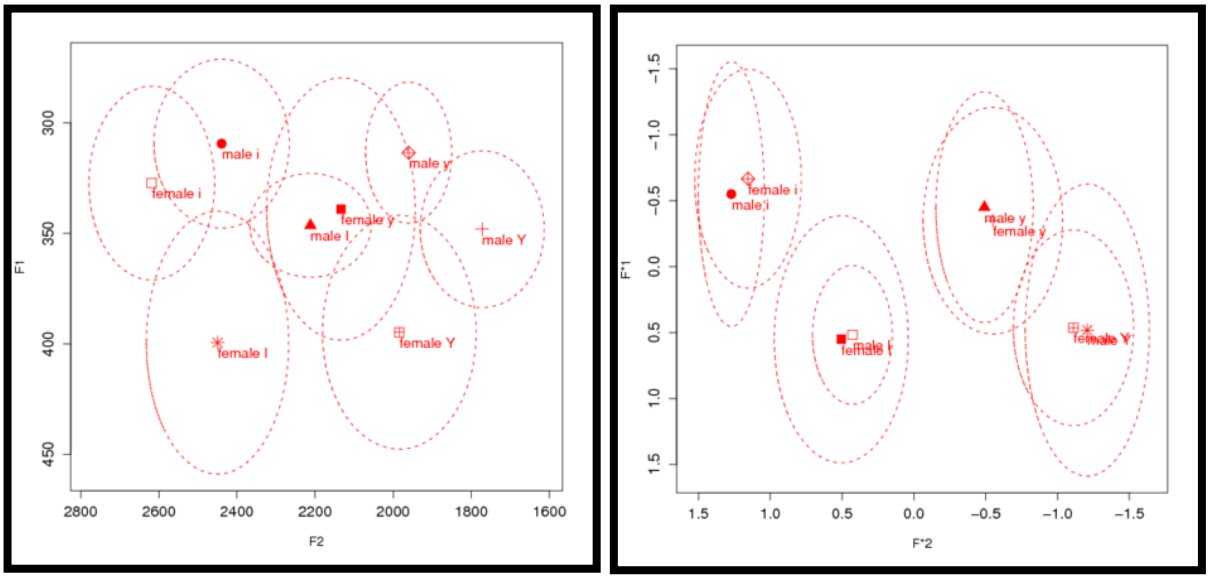
about the talker's regional background (or sociolinguistic information), and (c) minimize anatomical/physiological variation in acoustic representations of vowels" (2004b: 3099) and found that the Lobanov method was highly favourable. Since the three factors summed up by Adank *et al.* (2004b) are particularly important for the present study, the Lobanov method was opted for. An example of the effectiveness of ruling out sex-based differences by the Lobanov method can be seen in Figures 6.9 to 6.16, where the non-normalised results are juxtaposed to the normalised results for each of the two tasks of the two speaker groups.[25] Especially when the respective standard deviations (indicated by ellipses) are taken into account, we can see why normalising the results is necessary for a reliable representation. Of course, we have to be careful due to the limited male data, but where non-normalised male and female values sometimes lie outside of each other's standard deviation range, the normalised results show that this can often be ascribed to physiological differences based on sex (i.e. the different size of the vocal tracts).[26]

[25] In these figures (and the subsequent figures from the present section in which vowel plots are presented), /ɪ/, /ʏ/, /ʊ/ and /ʌ/ are represented here as capital I, capital Y, capital U and '^', respectively (cf. also Figure 5.5 in section 5.5.2). It should be noted that the NORM Vowel Normalization and Plotting Suite (Thomas & Kendall, 2007) used in the present study based the plot scales on the F1 and F2 input - caution is thus needed when interpreting scales that only represent one participant group.

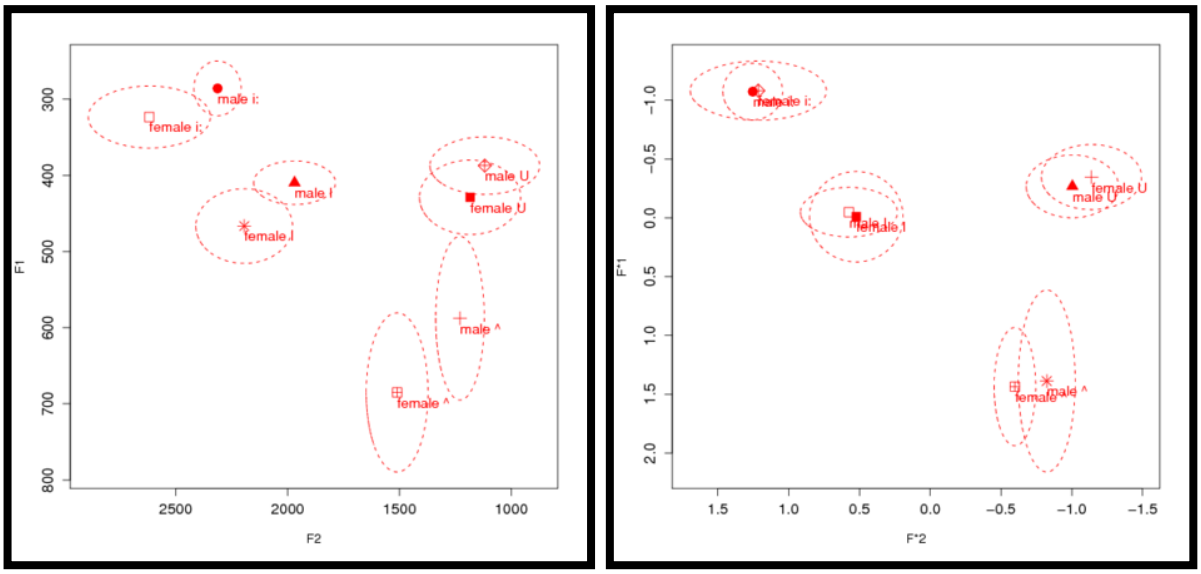
[26] It should be noted that the Lobanov method is a vowel-extrinsic method, which means that the vowels are viewed in relation with one another. Ideally, formant values for all vowels in the speakers' (L1 or L2) vowel system should be included; however, due to time limits, it was not feasible to measure productions of all Dutch and English vowels, when only four vowels of each vowel system are needed.



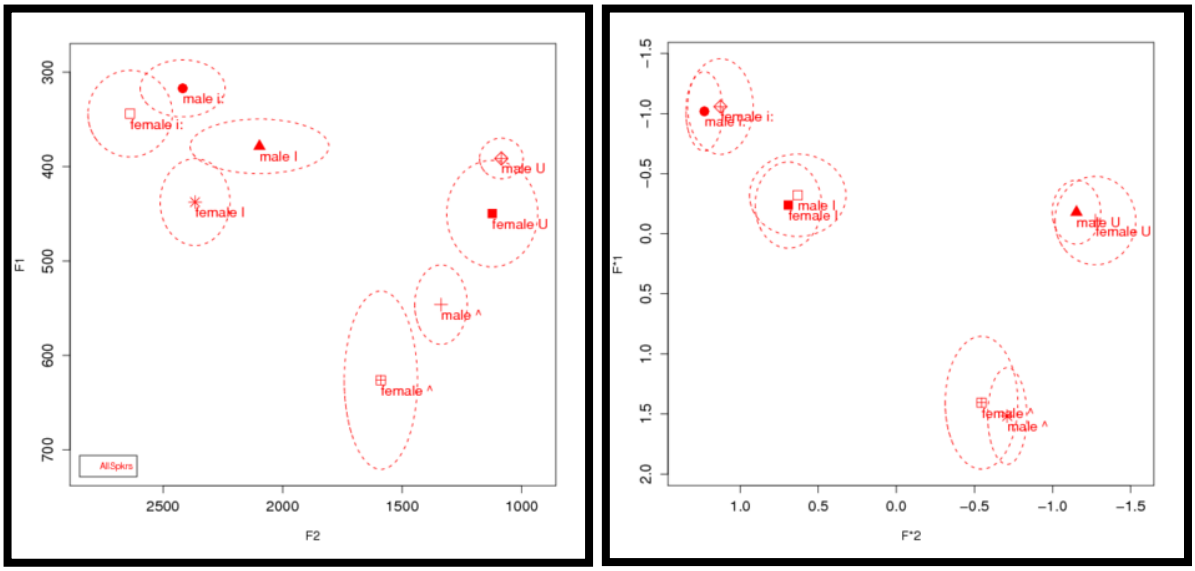
Figures 6.9 and 6.10. Ghent Dutch production task: non-normalised (left) and Lobanov normalised (right) results (the ellipses indicate the standard deviations).



Figures 6.11 and 6.12. Antwerp Dutch production task: non-normalised (left) and Lobanov normalised (right) results (the ellipses indicate the standard deviations).

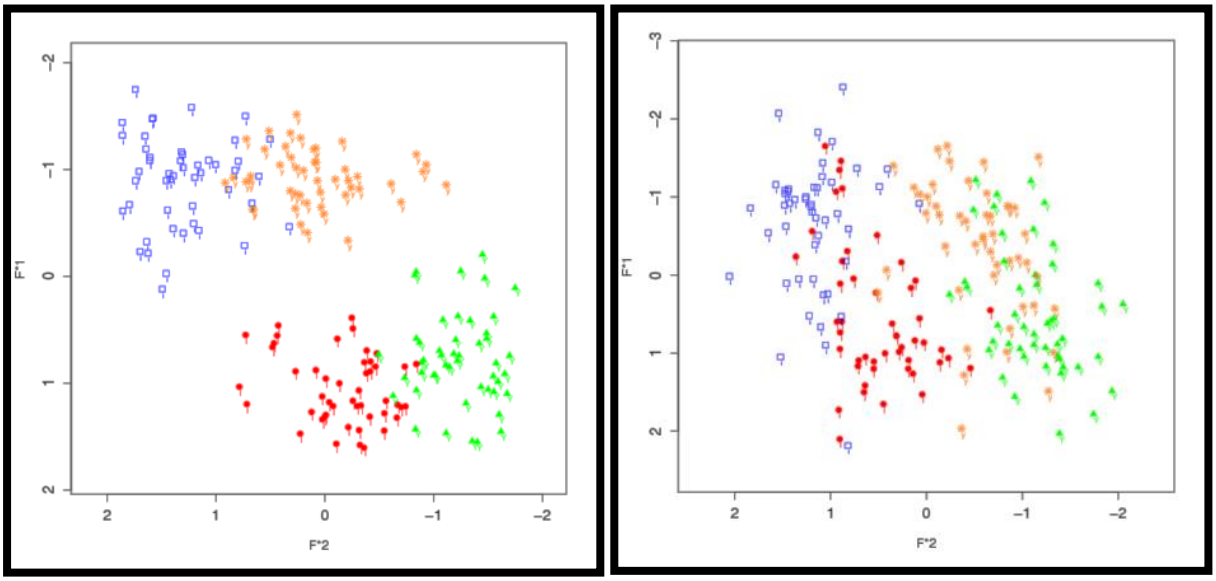


Figures 6.13 and 6.14. Ghent English production task: non-normalised (left) and Lobanov normalised (right) results (the ellipses indicate the standard deviations).



Figures 6.15 and 6.16. Antwerp English production task: non-normalised (left) and Lobanov normalised (right) results (the ellipses indicate the standard deviations).

A graphic representation of the normalised individual Dutch vowel productions per participant group is presented in Figures 6.17 and 6.18.



Figures 6.17 and 6.18. Dutch production task: Lobanov normalised individual productions of /i/ (blue), /ɪ/ (red), /y/ (yellow) and /ʏ/ (green) from the Ghent (left) and Antwerp (right) speakers.

Here, we can clearly discern differences between the two speaker groups. The predicted spectral overlap in Antwerp Dutch /i-ɪ/ and /y-ʏ/ productions, for example, immediately catches the eye - especially when compared to Ghent (East-Flemish) Dutch, where this overlap is nearly absent. The spectral differences are also apparent when looking at the statistical analyses presented in Table 6.5, in which the results for duration are also included, for the sake of completeness.

Table 6.5. P-values for the t-test for equality of means (Dutch target vowels). The significant results ($p \leq 0.05$) are indicated in green.

		F1 (p)	F2 (p)	duration (p)
Dutch /i/	male	0.247	0.001	< 0.001
	female	0.243	0.587	< 0.001
		F1 (p)	F2 (p)	duration (p)
Dutch /ɪ/	male	< 0.001	< 0.001	0.417
	female	< 0.001	< 0.001	0.232
		F1 (p)	F2 (p)	duration (p)
Dutch /y/	male	0.078	0.681	0.382
	female	0.042	0.917	0.184
		F1 (p)	F2 (p)	duration (p)
Dutch /ʏ/	male	< 0.001	0.001	0.066
	female	< 0.001	< 0.001	0.297

Combined with the Lobanov normalised visual representations of the spectral features of the vowel productions in Figures 6.17, 6.18 and 6.19, Table 6.5 generates a clear picture (keeping in mind the 5%-significance level): there is a highly significant durational difference ($p < 0.001$) between the Antwerp and Ghent productions of /i/ (due to the following /r/, this is not visible for /y/). At the same time, there is a highly significant spectral difference between the Antwerp and Ghent productions of /ɪ/ and /ʏ/ (except for the p-value of male F2 productions, which is $0.001 < 0.05$, all p-values for spectral features are < 0.001 for these vowels). Figures 6.17 and 6.18 show that Antwerp /ɪ/ bears considerable spectral overlap with Antwerp /i/. Analogously, Antwerp /ʏ/ bears spectral overlap with Antwerp /y/. These results provide a solid basis for accepting with high probability the fact that the predicted dialectal differences are very apparent in our sample of Ghent and Antwerp speakers.

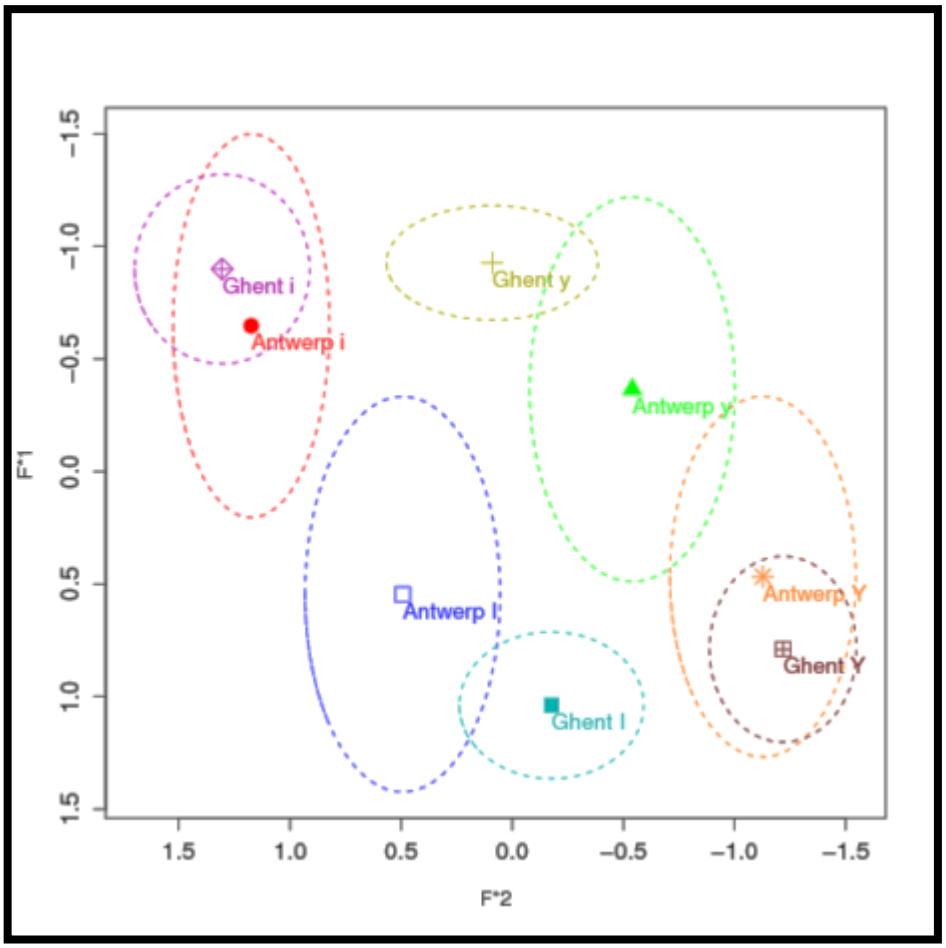
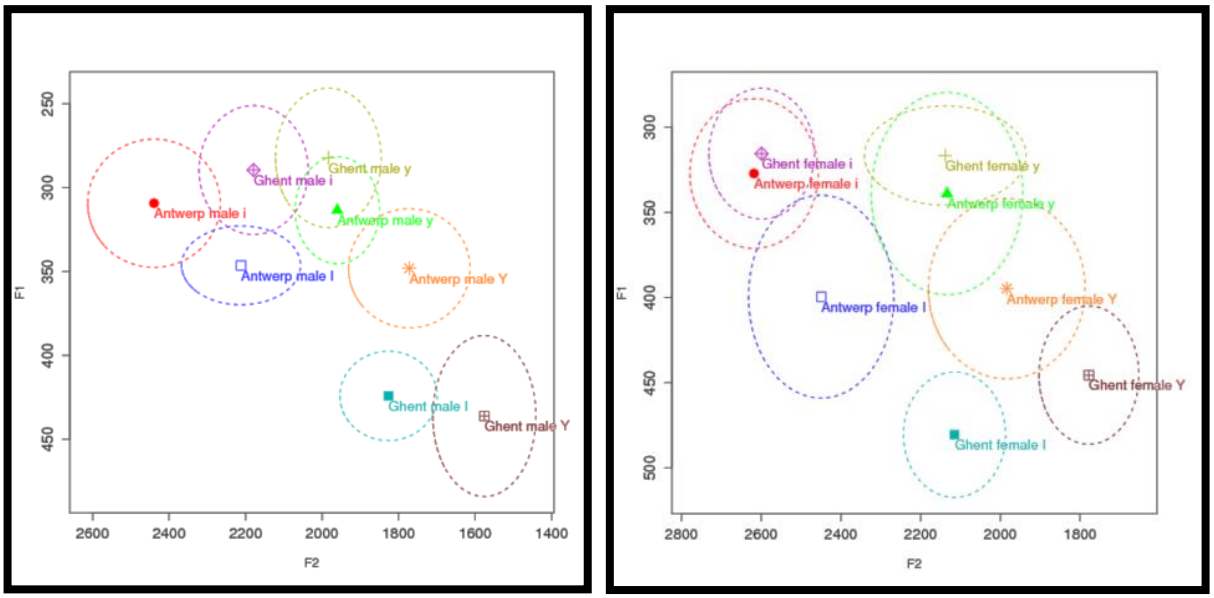


Figure 6.19. Dutch production task: Lobanov normalised average production of /i/, /ɪ/, /y/ and /ʏ/ from the Ghent and Antwerp speakers (the ellipses indicate the standard deviations).

For clarity's sake (Lobanov normalisation uses a transformation of the F1-F2 scale), the non-normalised average formant values are presented in Figures 6.20 and 6.21 (here, the male and female data are separated to prevent distorted results). Again, we can see very clearly the same results as presented above. Not only is the spectral relationship between the vowels of the two dialects different, the absolute F1 and F2 values of the same phonemes in the two dialects differ to a significant extent as well. Furthermore, we see in Figures 6.19, 6.20 and 6.21 that not only are short /ɪ/ and /ʏ/ (on average) produced more raised and fronted in Antwerp (Brabantine) Dutch, Antwerp long(er) /i/ and /y/ also have a slightly more open character than their Ghent equivalents, causing an even smaller spectral distance between these two vowel pairs in Antwerp dialect.



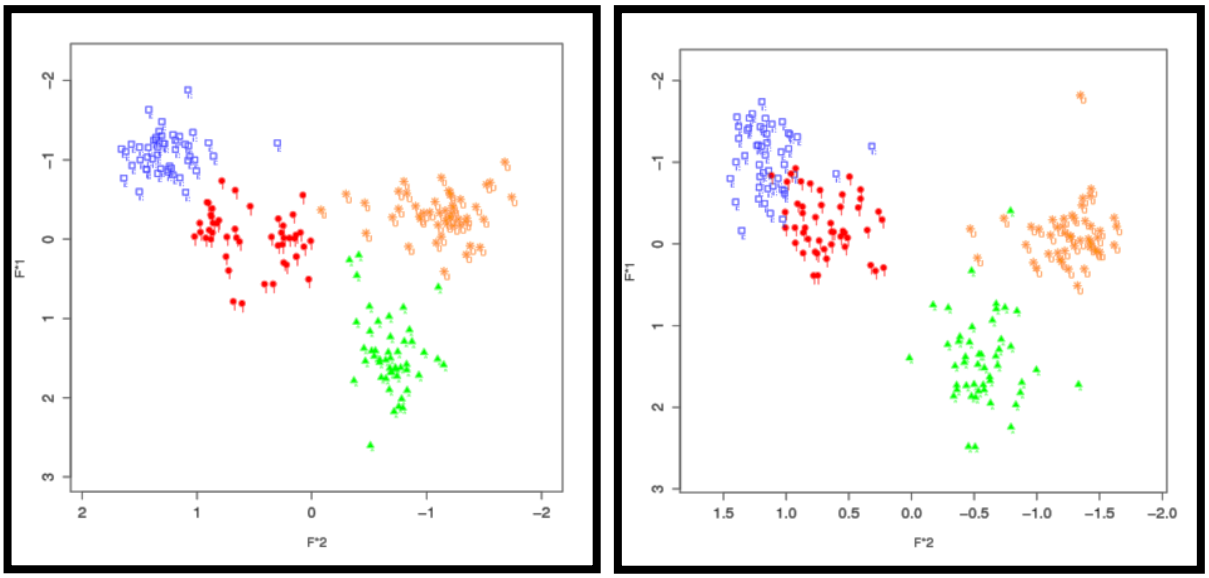
Figures 6.20 and 6.21. Dutch production task: average non-normalised Ghent and Antwerp male (left) and female (right) productions of /i/, /ɪ/, /y/ and /ʏ/ (the ellipses indicate the standard deviations).

For the English productions, a graphic representation of the normalised individual results per participant group is presented in Figures 6.22 and 6.23 below. Although the L2 differences are more subtle than the L1 differences, we can still see that Antwerp speakers, in comparison with Ghent speakers, have a smaller spectral contrast between their /i:/ and /ɪ/ productions, which corresponds to the predictions. The spectral differences in the /i:-ɪ/ contrast produced by Antwerp and Ghent speakers are also apparent when looking at the statistical analyses presented in Table 6.6 (again, the results for duration are also included, for the sake of completeness). English /ʌ/ also shows significant cross-dialectal differences. Moreover, when comparing Table 6.6 with Table 6.5, it becomes apparent that there are indeed fewer differences between the samples of speakers from the two dialectal regions in their L2 production than in their L1 production.

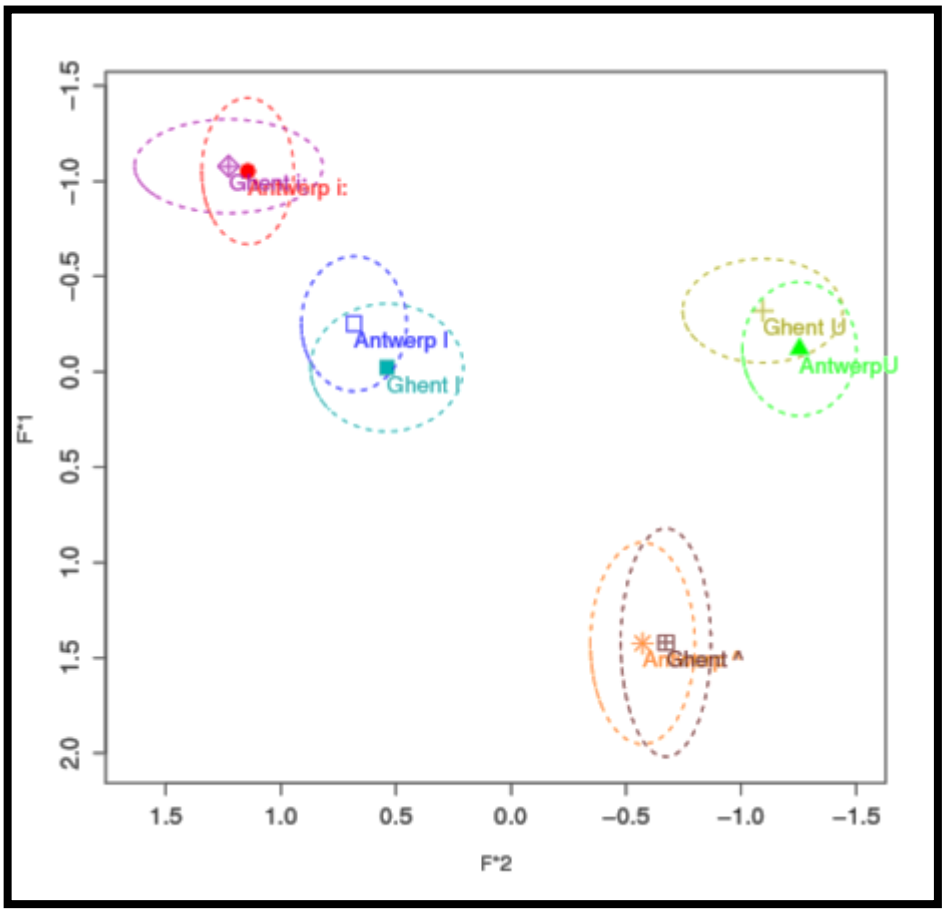
Table 6.6. P-values for the t-test for Equality of Means (English target vowels). The significant results ($p \leq 0.050$) are indicated in green.

		F1 (p)	F2 (p)	duration (p)
English /i:/	male	0.046	0.158	0.073
	female	0.049	0.559	0.240
		F1 (p)	F2 (p)	duration (p)
English /ɪ/	male	0.157	0.200	0.241
	female	0.010	< 0.001	0.026
		F1 (p)	F2 (p)	duration (p)
English /ʊ/	male	0.771	0.717	0.417
	female	0.099	0.220	0.911
		F1 (p)	F2 (p)	duration (p)
English /ʌ/	male	0.040	0.029	0.287
	female	0.015	0.027	0.614

In combination with the Lobanov normalised visual representations of the individual spectral vowel productions in Figures 6.22 and 6.23 below, Table 6.6 confirms our findings: indeed, the differences between Ghent and Antwerp English productions are less significant than the differences between the Dutch productions of the participants. The short vowels /ɪ/ and /ʌ/ show significant differences on the spectral level, and the female durational features of /ɪ/ are also significantly different (which is supported by Figures 6.7 and 6.8 in section 6.4.2.1). The significantly different F1 and F2 values of Antwerp and Ghent /ʌ/ are not expected. This could, however, be due to the limited participant sample and the subsequently greater impact of potential outliers. The normalised representation in Figure 24 neutralises this difference, but the non-normalised data in Figures 6.25 and 6.26 reveal that Antwerp /ʌ/ is produced more closed and fronted, on average. The overall results suggest that there are at least subtle dialectal differences in L2 production in our sample of Ghent and Antwerp speakers.



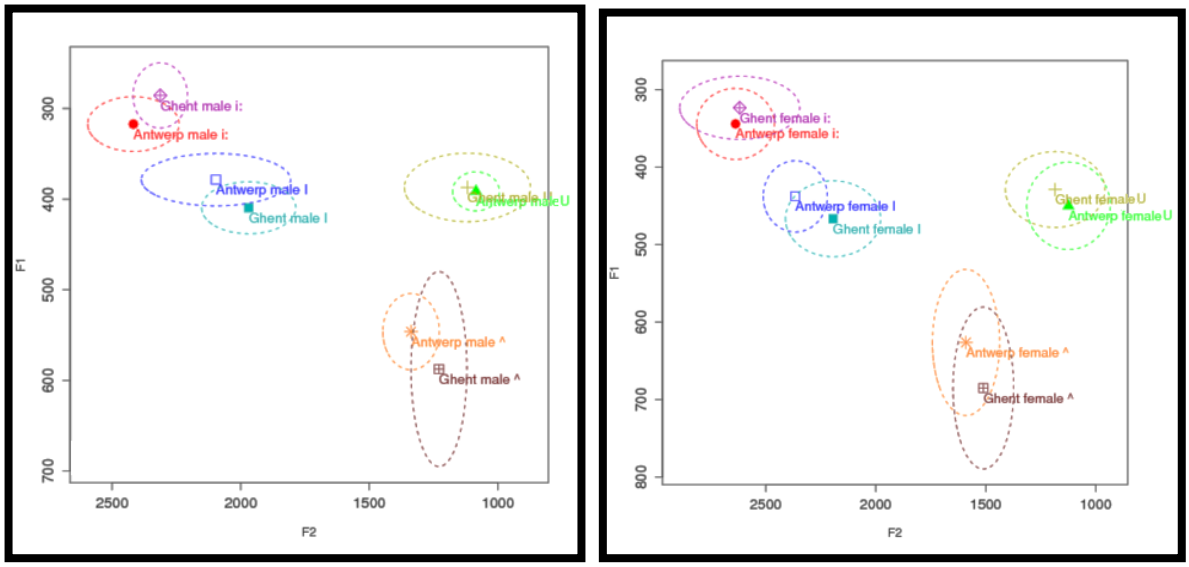
Figures 6.22 and 6.23. English production task: Lobanov normalised individual productions of /iː/ (blue), /ɪ/ (red), /ʊ/ (yellow) and /ʌ/ (green) from the Ghent (left) and Antwerp (right) speakers.



Figures 6.24. English production task: Lobanov normalised average production of /iː/, /ɪ/, /ʊ/ and /ʌ/ from the Ghent and Antwerp speakers.

Figures 6.25 and 6.26 present the average non-normalised values of the target vowels of the English production task, revealing the same results as presented above. Although

subtle (especially in comparison with the L1 productions), indications for a dialectal L1 influence on the L2 English productions are also visible in the absolute formant values. Figures 6.25 and 6.26 show, for instance, that the spectral differences between Ghent and Antwerp speakers with regard to the /i-ɪ/ are also transferred to the equivalent English /i:-ɪ/ contrast (be it to a lesser extent).



Figures 6.25 and 6.26. English production task: average non-normalised Ghent and Antwerp male (left) and female (right) productions of /i:/, /ɪ/, /ʊ/ and /ʌ/ (the ellipses indicate the standard deviations).

Section 6.5 compares the production results of the Dutch and English production tasks to studies on the vowels of Standard Dutch (Verhoeven & Van Bael, 2002a&b) and RP English (Hawkins & Midgley, 2005), and attempts to provide an explanation with the help of the theoretical models from the literature review and the results obtained in the perception experiment.

6.5 Discussion

6.5.1 Dutch production task

The results of the Dutch production task revealed that the predicted regional differences in L1 production can indeed be observed in the sample of Ghent and Antwerp speakers. To assess which participant group most closely approximated the Standard Dutch productions, the results from the present study are compared to the results from Verhoeven & Van Bael (2002b), in which (regional varieties of) Standard Dutch productions were measured.

With regard to the durational component, Figures 6.27 and 6.28 compare the duration averages for East-Flemish and Antwerp Standard Dutch (SD) from Verhoeven & Van Bael (2002b) with those obtained in the present tasks for the /i-ɪ/ (Figure 27) and /y-ʏ/ (Figure 28) contrasts.[27]

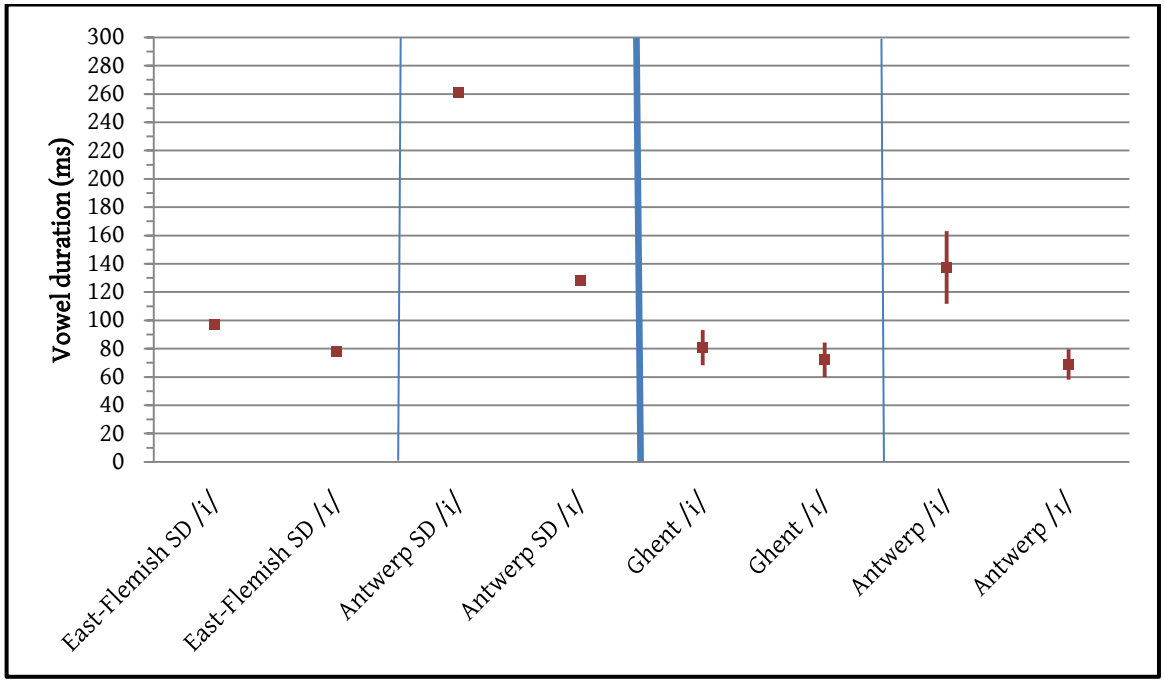


Figure 6.27. Average vowel durations from East-Flemish and Antwerp Standard Dutch /i-ɪ/ results from Verhoeven & Van Bael (2002b) (left) and from Ghent and Antwerp /i-ɪ/ results from the present study (right).

[27] Only the average vowel durations and formant values were presented in Van Bael & Verhoeven (2002b), and personal communication with the authors revealed that the individual results were no longer available, which is why presenting the standard deviations was not possible. For the sake of completeness, the results from Verhoeven & Van Bael that are relevant to the present study are presented in Appendix K.

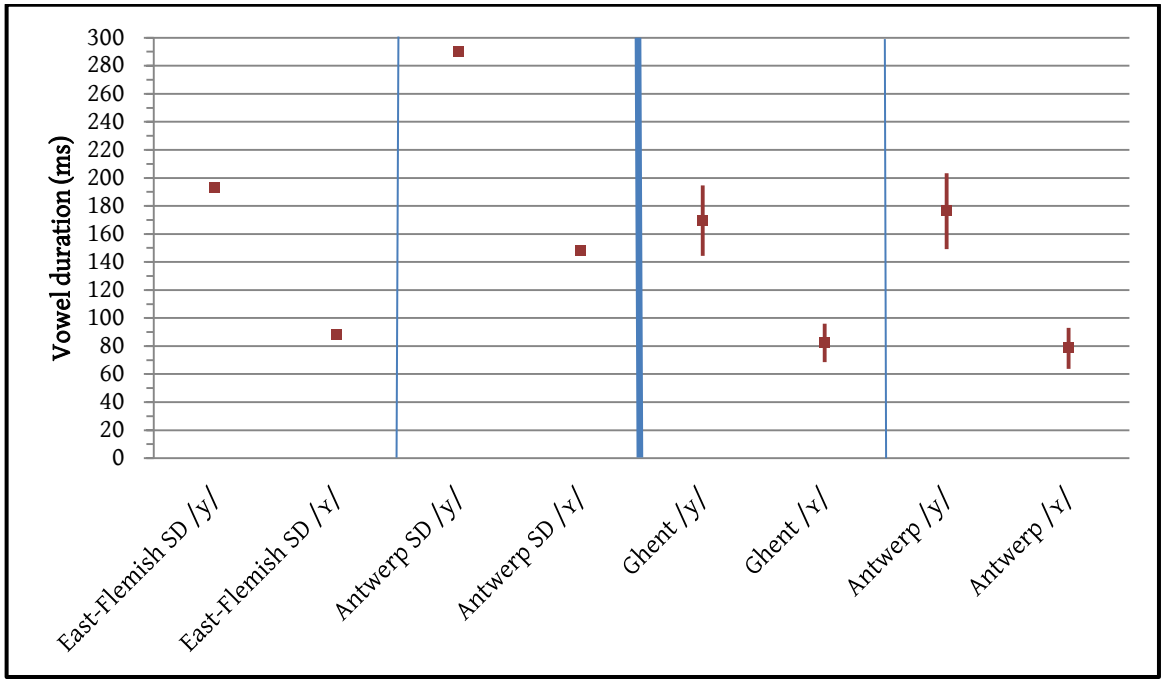


Figure 6.28. Average vowel durations from East-Flemish and Antwerp /y-ʏ/ results from Verhoeven & Van Bael (2002b) (left) and from Ghent and Antwerp /i-ɪ/ results from the present study (right).

Figures 6.27 and 6.28 reveal that the regional durational differences reported by Verhoeven & Van Bael (2002b) in their data are also apparent in the present speaker samples. Proportionally, the contrasts between the vowels are comparable in the two studies. That the absolute durational values of the vowels in the two pairs differ across the studies could be due to the different task set-up, or because a different age group was used in the two experiments (the average age of the participants from the Antwerp and East-Flemish group in Verhoeven & Van Bael, 2002a&b was 56.5 years). Nuyts furthermore argues that determination of duration is relative: "a sound is long/short relatively with regard to another sound in the same language" (1989: 25, my translation)[28]. Hence, it is perfectly possible for between-group differences in absolute vowel duration to appear, as long as the *relative* distinction between two vowels stays the same, as is the case here. That the duration of /y/ is almost the same for Ghent and Antwerp speakers in the present study, but not in Verhoeven & Van Bael, could have to do with the non-word /y/ stimuli

[28] "een klank is lang/kort relatief t.o.v. een andere klank in dezelfde taal."

used in the latter study: 'puut' and 'luus' (2002b: 6). Without the general lengthening from the /r/ consonant, the results from the present study perhaps would have shown a greater cross-dialectal contrast.

Our predictions that the Ghent data would correspond more to the East-Flemish data from Verhoeven & Van Bael (2002b) (and thus, as inferred in section 4.1.2, to the Standard Dutch norm), is confirmed by the present experiment. The Antwerp durational data from the present study, on the other hand, correspond (as expected) to the Antwerp production data from Verhoeven & Van Bael (2002b). In conclusion, we can thus state that there is indeed an L1 durational difference between Antwerp and Ghent speakers - in any case for the /i-ɪ/ contrast, the contrast which has proven, in the present study, to differ most between the two regional varieties (both on the durational as on the spectral level).

Concerning the spectral features of Standard Dutch, data with which to compare the results from the present study are very scarce. Verhoeven acknowledges that there are not many studies which have measured formant frequencies and, moreover, without the individual formant values of the speakers, and thus the possibility of vowel normalisation, comparison is difficult (personal communication, March 22, 2013). Despite the age differences between the participants of the two studies, the average formant frequencies provided in Verhoeven & Van Bael (2002b) are chosen for comparison. This is done for two reasons. First, the number of participants per speaker group is equivalent in both studies: Verhoeven & Van Bael (2002a&b) examine 12 Antwerp and 12 East-Flemish participants. Second, since Verhoeven & Van Bael (2002a&b) take into account the influence of sex on formant frequency values, their data averages are separated between men and women, so that they can be compared to the results of the present study. This separation causes normalisation (which would be impossible, since only averages are provided in Verhoeven & Van Bael, 2002b) to be of less importance, and makes comparison (at least of non-normalised formant values) possible.

For clarity's sake, the results of the comparison are again shown per vowel pair (this is possible because the non-normalised results do not require a vowel-extrinsic viewing of the vowels in relation to one another). Figures 6.29 and 6.30 present a spectral comparison of the male and female results for /i-ɪ/ from both studies; Figures 6.31 and 6.32 present a spectral comparison of the male and female results for /y-ʏ/ from both studies.

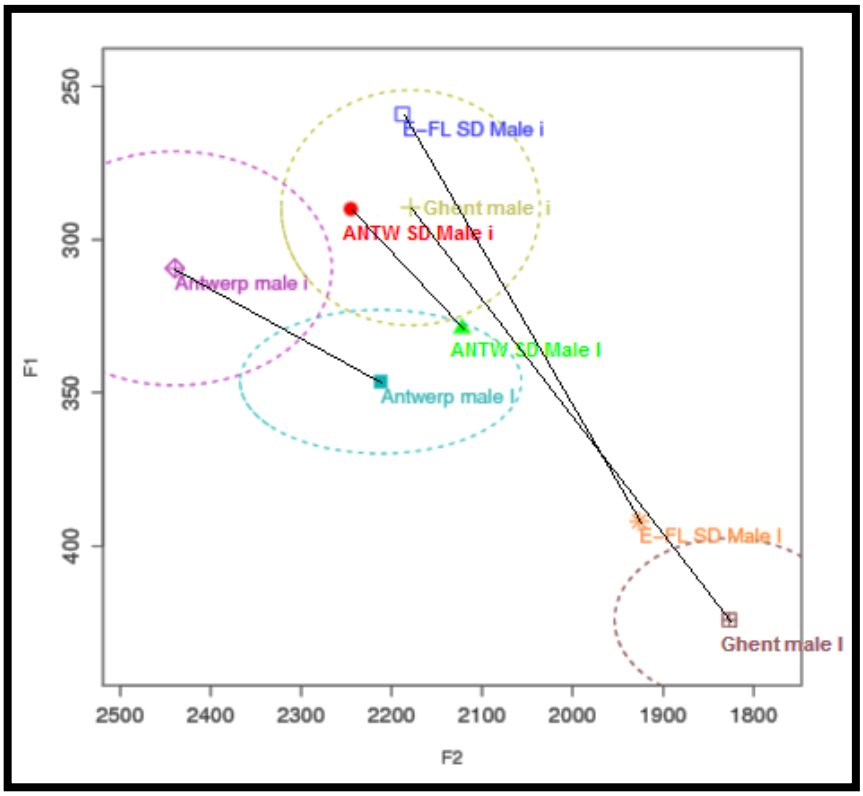


Figure 6.29. Comparison of the spectral features of the Antwerp and Ghent male /i-ɪ/ contrast from the present study (with indication of the standard deviation) with the ANTW(erp) and E(ast)-FL(anders) Standard Dutch results from Verhoeven & Van Bael (2002b). Vowel pairs produced by the same speaker group are connected through a full line to facilitate comparison.

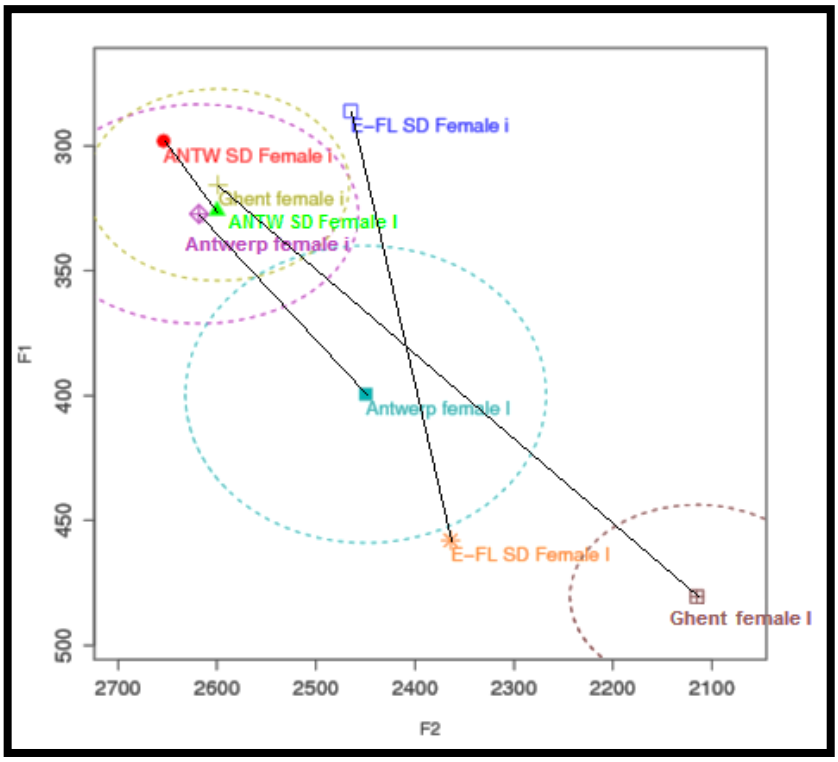


Figure 6.30. Comparison of the spectral features of the Antwerp and Ghent female /i-ɪ/ contrast from the present study (with indication of the standard deviation) with the ANTW(erp) and E(ast)-FL(anders) Standard Dutch results from Verhoeven & Van Bael (2002b). Vowel pairs produced by the same speaker group are connected through a full line to facilitate comparison.

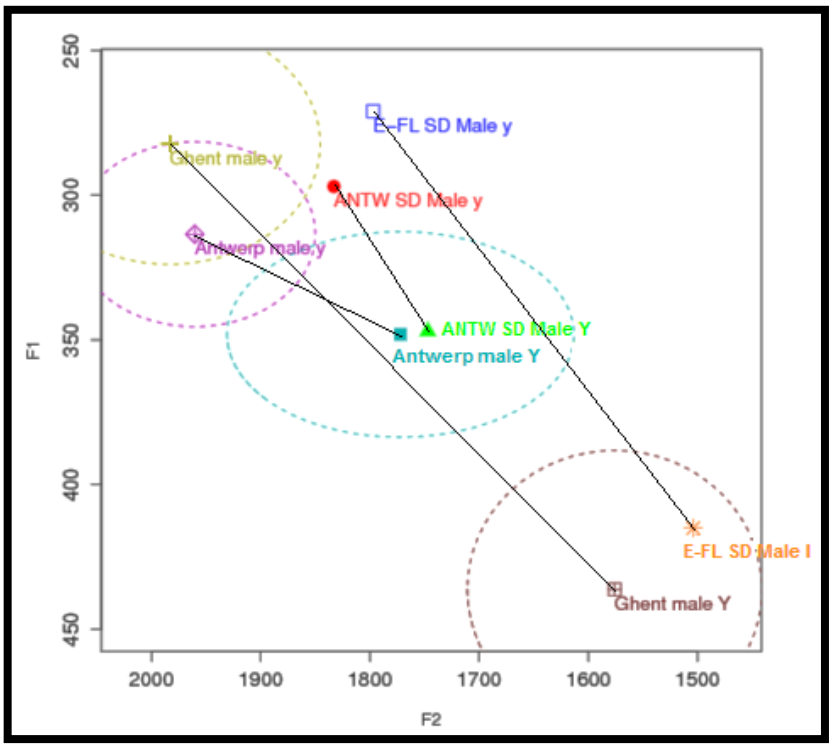


Figure 6.31. Comparison of the spectral features of the Antwerp and Ghent male /y-ʏ/ contrast from the present study (with indication of the standard deviation) with the ANTW(erp) and E(ast)-FL(anders) Standard Dutch results from Verhoeven & Van Bael (2002b). Vowel pairs produced by the same speaker group are connected through a full line to facilitate comparison.

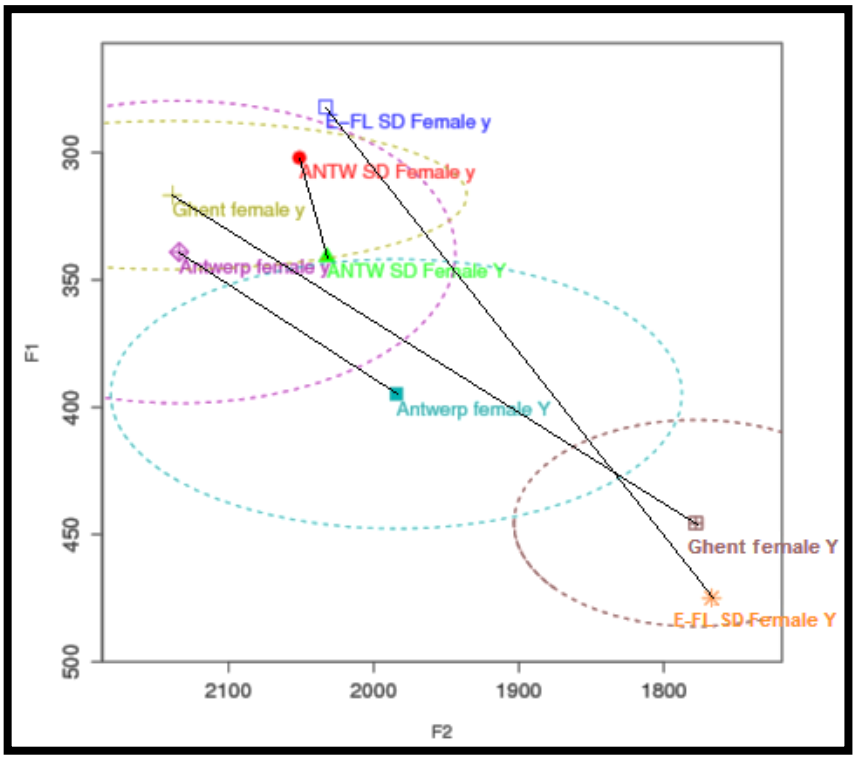


Figure 6.32. Comparison of the spectral features of the Antwerp and Ghent female /y-ʏ/ contrast from the present study (with indication of the standard deviation) with the ANTW(erp) and E(ast)-FL(anders) Standard Dutch results from Verhoeven & Van Bael (2002b). Vowel pairs produced by the same speaker group are connected through a full line to facilitate comparison.

The fact that the spectral distance between the Antwerp contrasts from Verhoeven & Van Bael (2002b) is generally smaller than the spectral distance between the contrasts from the present study could be due to the fact that the participants from the present study were younger, highly educated and predominantly female. These three factors generally predict a standard language-like direction of production (or at least a production that is more standard-like compared with that of the much older participants from Verhoeven & Van Bael, 2002b). Nevertheless, two groups can still be distinguished on the basis of spectral distance between the two vowel pairs: one group consists of the East-Flemish speakers from Verhoeven & Van Bael (2002b), together with the Ghent speaker sample from the present study. A relatively greater spectral distance in the two vowel contrasts can be observed in these two groups. This corresponds to the results from the perception experiment, and to the hypothesis which proposed that Ghent speakers will have a more Standard Dutch-like

production (as was pointed out in section 4.1.2, the East-Flemish productions can be seen as indicative for the Standard Dutch production).

The other group consists of the Antwerp speaker group from Verhoeven & Van Bael (2002b) and the sample of Antwerp speakers from the present study. The spectral distance between the two vowel pairs is smaller in comparison with the regional varieties from the other group. Compared with the results for duration (except for /y-ʏ/, due to aforementioned lengthening effect of /r/, but cf. the results from Verhoeven & Van Bael, 2002a&b), this corresponds to the hypothesis that Antwerp speakers put more weight on durational features than on spectral features to distinguish between the Dutch /y-ʏ/ and /i-ɪ/ vowel pairs. In Escudero's L2LP (2005) terms, the cue constraints for the durational feature are more developed in Antwerp listeners with regard to these vowel pairs. Learning tasks for an L2 which distinguishes only spectrally between such vowel pairs would involve adapting the L1 perception grammar to incorporate distinguishing L2 spectral features that would contribute to optimal L2 perception. Following Escudero's L2LP model (2005), L2 to L1 influence is not expected, as the different perception grammars remain separate. However, if Flege (1995) and Best & Tyler (2007) are followed, and L2 influence on L1 perception and subsequent production remains possible, it is not unthinkable that learning the spectral distinction (next to the durational distinction) in English /iː-ɪ/ could contribute to a more accurate perception and subsequent production by Antwerp listeners of the spectrally contrasted Standard Dutch /i-ɪ/ contrast, which they have difficulty perceiving and producing (cf. section 4.1.3).

The above comparison shows that the regional distinctions mentioned in the relevant literature are also apparent in the sample of Antwerp and Ghent speakers from the present study. There are clear regional differences in the production of /i ɪ y ʏ/, and these differences suggest that the phonetic categories are constructed in different ways for speakers from the two regional backgrounds: whereas the Antwerp speakers put more

emphasis on durational features to distinguish between the vowels of the two vowel pairs under consideration, the Ghent speakers rely more on spectral differentiation to accurately contrast the two vowel pairs (and by doing so, are more in accordance with the Standard Dutch production).

As was asserted in the three theoretical frameworks presented in section 2, L1 background plays an important role in L2 perception (and subsequent production). That there are indeed regional differences in the perception and production of L1 sounds by the two speaker groups from the present study, leads us to assume (as has already been shown in the presentation of the results in section 6.4.2.2) that influence on L2 production is possible. A comparison with RP English in section 6.5.2 will back up this claim.

6.5.2 English production task

The study of Hawkins & Midgley (2005) with which the results from the present study are compared, does not include durational measurements. It seems disadvantageous to compare the durational results from the present study to one RP English study, and the spectral results to another RP English study. Therefore, the durational component in this discussion will be dealt with by comparing the English durational results from the samples of Ghent and Antwerp speakers to each other in the light of their respective Dutch durational results. If we do so, it becomes obvious that the Ghent speakers are almost as successful as Antwerp speakers in the production of the English /iː-ɪ/ contrast, despite their not making a durational distinction between the relevant vowels in Dutch (cf. Figures 6.5, 6.6 and 6.7, 6.8). Regional differences for this contrast only manifest themselves subtly, in the durational distance between /iː-ɪ/, which is slightly larger in Antwerp speakers.

The Antwerp speaker sample, moreover, produces English /ʊ/ with the same duration as the Ghent speakers, despite the Dutch /u/ phoneme (to which English /ʊ/ is frequently

perceptually assimilated, cf. section 5.4.2) from Antwerp dialect being much longer than its East-Flemish counterpart.

These two observations provide evidence for the fact that the phonetic categories for the English vowels /iː ɪ ʊ ʌ/ are successfully constructed by both groups (or in Escudero's 2005 L2LP terms, that the cue constraints were successfully altered to optimally perceive and produce the L2), at least concerning the durational features. The theoretical models indeed leave space for accurate, native-like L2 perception (and subsequent production). In line with Escudero's L2LP (2005), one could say that the learning tasks for the durational features of the vowels in the two participant groups from the present study have been successfully completed by both regionally differing speaker samples (especially when comparing the results to the L1 results for duration, and when comparing the L2 results to the durational results of the native British RP English speaker in Figure 5.4 in section 5.5.2).

To examine which speaker group's spectral results most closely resemble those of RP English, the results from the present study are compared to those from Hawkins & Midgley (2005), in which formant frequencies of RP English vowels from people in different age groups were measured. This study is especially convenient because, unlike the studies that measure Dutch formant frequencies, the authors provide an appendix with the individual formant frequency values per speaker and per target stimulus. Moreover, one of the age groups under consideration consists of speakers aged 20 to 25, which is approximately the age group to which the participants from the present study belong. Although the target stimuli were inserted in an identical /hVd/ consonantal context, they were meaningful, as is the case for the present study.[29] The fact that the RP speaking participants were all male does not have to be a problem, since the results can be normalised. For the sake of

[29] The target stimuli for /iː ɪ ʊ ʌ/ in Hawkins & Midgley (2005) were 'heed' /hiːd/, 'hid' /hɪd/, 'hood' /hʊd/ and 'hud' /hʌd/.

completeness, the formant values from Hawkins & Midgley (2005) relevant to the analysis in the present subsection are presented in Appendix L.

As vowel-extrinsic Lobanov normalisation requires as many vowels as possible to be included in the analysis, the results will not be separated into the two contrasts, as was the case in the previous subsection. Figure 6.33, then, presents the Lobanov normalised spectral results for /iː-ɪ/, /ʊ/ and /ʌ/ for the English productions of the Antwerp and Ghent speaker groups of the present study, compared with those of the RP English speaker group with ages ranging from 20 to 25 years.

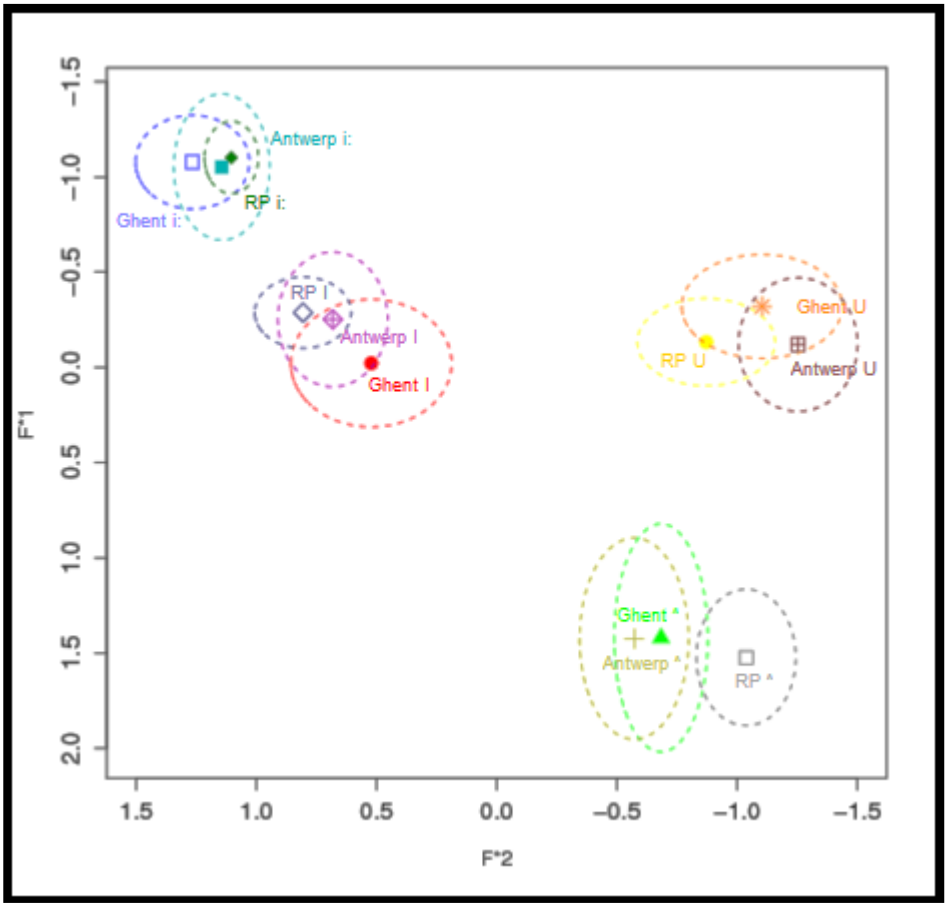


Figure 6.33. Comparison of the English /iː ɪ ʊ ʌ/ productions of the Ghent and Antwerp speaker samples from the present study to RP results for the same vowels from 5 speakers aged 20 to 25, taken from Hawkins & Midgley (2005). The standard deviations are indicated by ellipses.

The graph reveals that the English L2 productions of the Ghent and Antwerp participants are much more similar, compared to their L1 productions. Accurate L2

production seems almost attained by the two speaker group samples. However, some minor differences still occur, cross-dialectally and cross-linguistically.

Cross-dialectally, the members of the Antwerp /iː-ɪ/ contrast are indeed spectrally closer to each other than those of the Ghent /iː-ɪ/ contrast. This could potentially be due to transfer from the L1 regional variety. However, what is striking is that when compared to RP, the Antwerp spectral qualities seem to contribute to a production that is more similar to that of modern RP. This was not predicted by the hypotheses; on the contrary, Ghent English productions seem to suffer (comparatively) from the greater spectral distinction in their L1 dialect.

Cross-linguistically, with regard to English /ʊ/, the results seem to indicate that the hypothesis that this vowel would be perceptually assimilated to Dutch /u/ can be confirmed; both Ghent and Antwerp /ʊ/ productions are more backed than RP /ʊ/. Cross-dialectally, Antwerp /ʊ/ is produced more backed and open than Ghent /ʊ/, which appears to be in line with our hypothesis that Antwerp /ʊ/ could be produced further away from the vowel space in which Antwerp listeners have already constructed two phonetic categories (i.e. for the more or less spectrally coinciding /y-ʏ)/

The productions of /ʌ/, on the other hand, are more fronted both in the Ghent and Antwerp productions, in comparison with the RP productions. The /ʌ/ productions of the two speaker groups do not seem to be too deviating from each other, especially when compared to the RP productions of /ʌ/. The frontedness and the range of standard deviations in the L2 /ʌ/ productions correspond to the image in the perception tasks, and to the predictions with regard to the perceptual assimilation and subsequent production of English /ʌ/: if the vowel was not produced accuratey, it would either be assimilated to the more closed Dutch /ʏ/ or to the more open Dutch /ɑ/ and /aː/. Indeed, when we look at how the relevant vowels are situated on the IPA chart (cf. Figure 6.34), we get an image that

is very similar to that in Figure 6.33: /ʏ/ and /aː/ are indeed produced more fronted, and /ʏ/ is produced closer while /ɑ/ and /aː/ are produced more open. This direction of production is indeed apparent in Figure 6.33, and could be explained by perceptual assimilation and subsequent production (cf. also the results from the perception tasks in section 5.4.2).

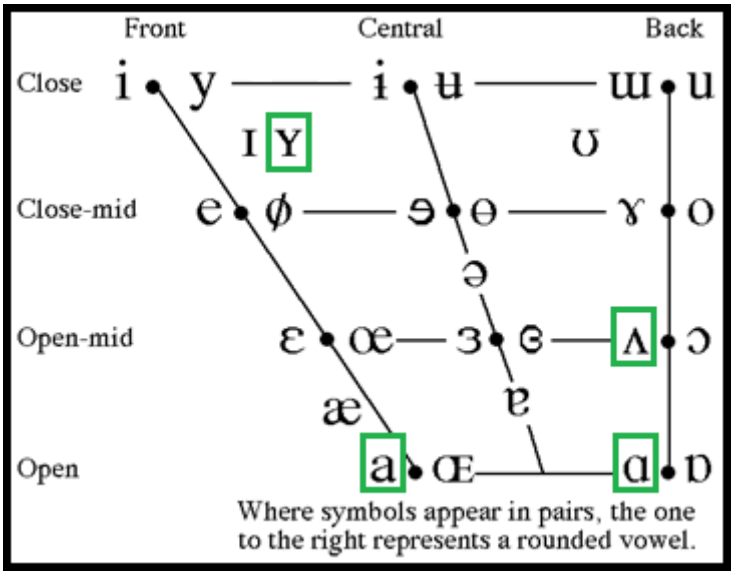


Figure 6.34. IPA vowel chart. The relevant vowels are indicated in green (adapted from http://www.phonetics.ucla.edu/course/chapter1/vowels.html).

It should be pointed out that, in the present comparison, caution is still needed. Of course, RP English is a high prestige variant, and inter-speaker variability would not be expected to such a great extent as in L2 English production (which can also be seen in the generally smaller standard deviations for RP productions in Figure 6.33). Nevertheless, 5 participants is still a rather limited sample, and outliers could influence the outcome. Nevertheless, L2 regional differences still seem to occur between the samples of Antwerp and Ghent English speakers, and these differences seem to be at least partly based on how the L1 vowels relate to one another in the specific L1 regional varieties, providing evidence for the fact that those L1 dialects can influence L2 production. However, the L2 differences seem rather small, proving that accurate L2 perception and production is possible for both speaker groups, despite the considerable differences in their L1 productions. That a

contrast which deviates from the L1 standard language need not be disadvantageous in the L2 production, is proven by the Antwerp English /iː-ɪ/ production as illustrated in Figure 6.33, where the closer spectral contrast between these two vowels contributes to a more RP-like production.

7 Conclusions and implications

The present study was incited by recent developments in the area of L2 phonology that underlined the importance of taking into account L1 regional variation when examining L2 perception and production. The study has shed light on two issues: the influence of L1 regional vowel differences on the perception and production of the L1 standard variety (discussed in section 7.1), and its influence on L2 perception and production (section 7.2).

7.1 L1 regional variation: impact on L1 vowel perception and production

A first aim of this study was to examine whether the regional L1 differences between Ghent East-Flemish Dutch and Antwerp Brabantine Dutch /i-ɪ/ and /y-ʏ/ as described in the literature (reviewed in section 4) could be observed in new samples of 12 Ghent and 12 Antwerp young adult participants.

With regard to L1 Dutch perception, the categorisation distributions from the two participant groups confirmed the predicted dialectal differences. The more accurate performance by Ghent listeners in categorising all four Standard Dutch stimuli (/i ɪ y ʏ/) confirms that, with respect to these vowels, Ghent (East-Flemish) Dutch is more similar to Standard Dutch than Antwerp (Brabantine) Dutch. The Antwerp participants' comparatively lower performance on the categorisation of Standard Dutch stimuli furthermore revealed a cue weighting in Antwerp perception grammar that leans more to durational than to spectral features (at least to distinguish between the vowel pairs examined in the present study).

The picture sketched by the Dutch perception task was confirmed in the Dutch production task. A greater durational and more limited spectral distinction in Antwerp Dutch /i/ and /ɪ/ productions was revealed. These productions differed significantly from Ghent Dutch /i-ɪ/ productions, with their greater spectral and durationally more limited

distinctions. For /y-ʏ/, no significant durational production differences between the two groups were found (but cf. section 6.4.2.2), yet the spectrally significantly differing /ʏ/ productions showed a pattern which was similar to that of the /i-ɪ/ contrast. Everything considered, the two Dutch tasks thus illustrated that the boundaries of the phonological categories for /i-ɪ/ and /y-ʏ/, and the weight attached to phonetic features to distinguish between these two vowel pairs were significantly different in Ghent (East-Flemish) and Antwerp (Brabantine) Dutch. This conclusion generates some interesting implications for both education and future research.

On an educational level, the observation that stable phonological categories that incorporate regional phonetic features have been constructed by the Antwerp participants could prove problematic, at least when Standard Dutch acquisition is intended. For instance, in 45.8% of the cases in the Dutch perception task, Antwerp listeners perceptually assimilate Standard Dutch /i/ to /ɪ/. Moreover, section 5.4.1 reveals that Antwerp listeners' perceptual performance is more inaccurate than Ghent listeners' perceptual performance for all four Standard Dutch stimuli. Standard language production was not aimed at (i.e., the participants were not explicitly asked to speak Standard Dutch, although the formal setting of the task was expected to elicit more standard-like productions), but the results from the present study still confirm this deviating featural pattern of vowel distribution. Further research that specifically focuses on Antwerp speakers' Standard Dutch productions could shed more light on the matter. If equivalence classification of Standard Dutch to *similar* Antwerp (Brabantine) Dutch vowels occurs, Flege's SLM (1995) predicts a great deal of difficulties in the acquisition of the equivalent Standard Dutch vowel contrasts (which is confirmed by the observations in section 4.1.3). Teachers could be encouraged to develop specific teaching strategies to properly handle pronunciation difficulties experienced by

students. Greater awareness of regionally deviating vowel qualities in students (and teachers[30]) could facilitate Standard Dutch perception and subsequent production.

Regarding the perception task, the present study was restricted to the perception of Standard Dutch vowels. In a region like Flanders, where dialectal variation is still very present (at least in comparison with surrounding countries, cf. Van de Velde, 2002: 137), it could be fruitful for future studies to adapt a cross-dialectal point of view to examine how listeners from different L1 regional backgrounds perceive vowels from other L1 regional backgrounds. Since the distinction between a dialect and a language is not always clearly delineated, and can even depend on the definition used[31], there are no theoretical reasons not to apply concepts such as perceptual assimilation (Best, 1995), equivalence classification (Flege, 1995) or perception grammars and cue constraints (Escudero, 2005) to cross-dialectal L1 perception and production studies as well.

The highly significant differences between Ghent and Antwerp L1 perception and production are especially striking when taking into account the nature of the participant groups. Most of the participants were female, and all of them were young and higher education students - three sociolinguistic parameters which would contribute to a standard language-like production and perception, and thus to small differences between the two regional groups. As shown in sections 5 and 6, this is not the case, and dialectal variation persists between the participant groups. Although caution is still needed, generalisations for the whole population are more justified in studies like the present one, that examine regional variation in samples in which dialectal influence is not expected to a great extent. Future studies could examine whether the dialectal differences are even greater in samples of, for instance, predominantly male, or older, or less well-educated participants.

[30] Although Standard Dutch is the educational norm, teachers rarely consistently follow this guideline (Delarue, 2011: 9-14).

[31] Cf. Bakkum (2009: 341-360) for an overview of the definitions (applied to the question of whether or not Faliscan can be regarded as a dialect of Latin).

7.2 L1 regional variation: role in L2 vowel perception and production

The second aim of this study was to examine if and how these regional L1 vowel differences influenced the L2 perception and production of English /iː-ɪ/, /ʊ/ and /ʌ/. The theoretical L2 perception and production frameworks presented in section 2 (which all emphasised the importance of the L1 production environment and vowel system configuration) suggested that such an influence was highly plausible, since Ghent (East-Flemish) and Antwerp (Brabantine) Dutch differ to a great extent regarding the studied vowels. The studies reviewed in section 3, where L1 regional influence on L2 vowel perception and production was revealed in many cases, further validated the prediction that Ghent and Antwerp L2 perception and production would differ.

In the cross-language perception task, Ghent participants were more 'successful'[32] than Antwerp participants in categorising English /iː/ and /ɪ/ as Dutch /i/ and /ɪ/, respectively. The differences were in line with the Dutch perception task results. English /ʊ/ and /ʌ/ are mapped significantly differently onto Dutch L1 vowel categories by the two participant groups, and the different distribution could be accounted for by regional L1 differences.

Antwerp participants nevertheless mapped RP English /iː/ onto Dutch /i/ to a far greater extent than Standard Dutch /i/ in the Dutch perception task. This is especially striking since the participants were unaware of the fact that the task contained English stimuli: according to Escudero's L2LP (2005), the L1 perception grammar (which is separate from the L2 perception grammar) was still activated; optimal perception for Antwerp listeners thus benefitted from the longer English /iː/ stimuli. This leads to a remarkable observation: if a vowel from an L1 regional variety bears more similarity to an L2 vowel (or

[32] 'Successful' in this sense should be viewed as 'according to what was expected for Dutch listeners when comparing how the L1 and L2 vowels generally relate to one another', and the expression here does not contain in itself a value judgment.

is weighted towards the same cues) than to the equivalent L1 standard language vowel, perception of the L2 vowel could prove to be more accurate than perception of the L1 standard language vowel.

Despite the categorisation distributions being roughly similar between the two participant groups, the Ghent and Antwerp results for English /ʊ/ and /ʌ/categorisations were significantly different, and these differences could be accounted for by the way in which the vowel distributions in the L1 regional variations were organised. That, even with regional differences, English /ʌ/ was categorised mainly as Dutch /a:/ and /ɑ/ by both participant groups indicates an instance of the SUBSET scenario from Escudero's L2LP (2005).

With regard to the English production task, regional L1 differences revealed to be of some influence on the L2 production of English vowels, yet the two participant groups differed to a lesser extent than would be expected when taking their L1 productions into consideration. What is especially striking is that the smaller spectral distinction between the vowels of the L1 /i-ɪ/ vowel pair produced by Antwerp speakers seems to contribute to a more native-like production of the RP English /i:-ɪ/ contrast than that produced by Ghent speakers.

Although minor regional differences in L2 production can be observed, the English productions of Ghent and Antwerp speakers are in general comparable. Similar productions by both speaker groups, despite extensive regional L1 differences, indicate that the L2 vowel contrasts have been successfully acquired - the participants from the present study could then be regarded as experienced L2 learners. Since the L1 production backgrounds for the two participant groups are expected to differ to a great extent, and this production background is of major importance to L2 perception and production, the logical conclusion would be that different L2 vowel learning tasks are used by the Antwerp and Ghent

participants, so that different L2 acquisition patterns arises (nevertheless generating a similar L2 performance). Future studies could adopt a developmental point of view to study how L2 acquisition is structured in L2 learners from different L1 dialectal backgrounds. We believe that the present study has convincingly shown that extensively differing L1 regional perception and production patterns do not always lead to extensively differing L2 perception and production patterns, since L1 regional influences may be overruled by the L2 learning process in sufficiently experienced L2 speaker-listeners.

8 References

Adank, Patti, Roeland van Hout & Roel Smits. 2004a. An acoustic description of the vowels of Northern and Southern Standard Dutch. *Journal of the Acoustical Society of America* 116 (3). 1729-1738.

Adank, Patti, Roel Smits & Roeland van Hout. 2004b. A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America* 116 (5). 3099-3107.

Bakkum, Gabriël Cornelis Leonides Maria. 2009. *The Latin dialect of the Ager Faliscus: 150 years of scholarship.* University of Amsterdam doctoral dissertation. Amsterdam: Vossiuspers UvA - Amsterdam University Press.

Belemans, Rob, Jan Goossens, Antonius Angelus Weijnen, Roeland Van Hout & Willy Van Langendonck. 2000. *Woordenboek van de Brabantse dialecten. 3: Inleiding; klankgeografie van de Brabantse dialecten.* Assen: Koninklijke Van Gorcum.

Best, Catherine T. 1995. A Direct Realist View of Cross-Language Speech Perception. In W. Strange (Ed.), *Speech Perception and linguistic experience: issues in cross-language research.* 171-204. Timonium, MD: York Press.

Best, Catherine T. & Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In M.J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production.* 13-34. Amsterdam: John Benjamins.

Boersma, Paul. 1998. *Functional Phonology*. University of Amsterdam doctoral dissertation. The Hague: Holland Academic Graphics.

Boersma, Paul & Kateřina Chládková. 2011. Asymmetries between speech perception and production reveal phonological structure. *The 17th International Congress of Phonetic Sciences: Hong Kong, China, August 17-21, 2011: congress proceedings*. 328-331. Hong Kong: Department of Chinese, Translation & Linguistics, City University of Hong Kong.

Boersma, Paul & David Weenink. 2013. Praat: doing phonetics by computer [computer program]. Version 5.3.45, retrieved 15 April 2013 from http://www.praat.org/.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22. 129-159.

Chládková, Kateřina & Václav Jonáš Podlipský. 2011. Native dialect matters: Perceptual assimilation of Dutch vowels by Czech listeners. *Journal of the Acoustical Society of America* 130 (4). 186-192.

Chládková, Kateřina & Silke Hamann. 2011. High vowels in Southern British English: /u/-fronting does not result in merger. *The 17th International Congress of Phonetic Sciences: Hong Kong, China, August 17-21, 2011: congress proceedings*. 476-479. Hong Kong: Department of Chinese, Translation & Linguistics, City University of Hong Kong.

Collins, Beverley & Inger M. Mees. 2003. *The phonetics of English and Dutch - Fifth Edition.* Leiden: Brill.

De Pelsmacker, Patrick & Patrick Van Kenhove. 2012. *Marktonderzoek - Methoden en toepassingen - Derde Editie.* Amsterdam: Pearson.

Debaene, Mathijs. 2012. *The effect of first language regional variation on second language vowel production. An experimental study on the production of English /i:/, /ɪ/ and /ɛ/ by Antwerp Brabantine Dutch speakers.* Ghent: University of Ghent BA research paper.

Delarue, Steven. 2011. Standaardtaal of tussentaal op school? De paradoxale dubbele kloof tussen taalbeleid en taalgebruik. *STUDIES VAN DE BKL* 6. 1-20.

Deterding, D.H. 1990. *Speaker Normalisation for Automatic Speech recognition*. University of Cambridge doctoral dissertation.

Escudero, Paola. 2005. *Linguistic perception and second language acquisition : Explaining the attainment of optimal phonological categorization.* University of Utrecht doctoral dissertation. Utrecht: LOT.

Escudero, Paola & Paul Boersma. 2004. Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26. 551-585.

Escudero, Paola, Ellen Simon & Holger Mitterer. 2012. The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception. *Journal of Phonetics* 40 (2). 280–288.

Escudero, Paola & Daniel Williams. 2012. Native dialect influences second-language vowel perception: Peruvian versus Iberian Spanish learners of Dutch. *Journal of the Acoustical Society of America* 131(5). 406-412.

Flege, James Emil. 1995. Second-language speech learning: theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and linguistic experience: issues in cross-language research*. 229-273. Timonium, MD: York Press.

Flege, James Emil. 1997. English vowel production by Dutch talkers: More evidence for the "new" vs. "similar" distinction. In A. James & J. Leather (Eds.), *Second-language speech, structure and process*. 11-52. Berlin: Mouton de Gruyter.

Flege, James Emil. 2005a (15 - 17 June). Evidence for plasticity in studies examining second language speech acquisition [conference presentation]. Presented at *ISCA Workshop on Plasticity in Speech Perception*. London, United Kingdom.

Flege, James Emil. 2005b (14 - 15 April). Origins and Development of the Speech Learning Model [conference presentation]. Presented at *1st ASA Workshop on L2 Speech Learning.* Vancouver, BC.

Flege, James Emil, Murray J. Munro & Ian R.A. MacKay. 1995. Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America* 97. 3125–3134.

Fowler, Carol A. & Lawrence D. Rosenblum. 1989. The perception of phonetic gestures. In I.G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the motor theory of speech perception*, 33-59. Hillsdale, NJ: Erlbaum.

Gramley, Vivian. 2010. Universität Bielefeld: Acoustic Phonetics [website]. http://www.uni-bielefeld.de/lili/personen/vgramley/teaching/HTHS/acoustic_2010.html. Last updated 9 November 2010. Consulted 31 May 2013.

Harrington, Jonathan, Felicitas Kleber & Ulrich Reubold. 2011. The contributions of the lips and the tongue to the diachronic fronting of high back vowels in Standard Southern British English. *Journal of the International Phonetic Association* 41 (2). 137-156.

Hawkins, Sarah & Jonathan Midgley. 2005. Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association* 35 (2). 183-199.

IBM Corp. 2012. IBM SPSS Statistics for Windows [computer program]. Version 21.0, retrieved 5 July 2013 from http://www.athenax.ugent.be/. Armonk, NY: IBM Corp.

Keulen, Ronny, Ton van de Wijngaard & Jacques Van Keymeulen. 2010. Regionale Woordenboeken [website]. https://fuzzy.arts.kuleuven.be/rewo/. Last updated April 2010. Consulted 27 May 2013.

Kluender, Keith R., Randy L. Diehl & Beverly A. Wright. 1988. Vowel-length differences before voiced and voiceless consonants: An auditory explanation. *Journal of Phonetics* 16 (2). 153-169.

Koopmans-van Beinum, F. 1980. *Vowel contrast reduction. An acoustic and perceptual study of Dutch vowels in various speech conditions*. University of Amsterdam doctoral dissertation. Amsterdam: Academische Pers.

Lehiste, I. 1988. *Lectures on Language Contact*. Cambridge, MA: MIT Press.

Leussen, Jan-Willem van, Daniel Williams & Paola Escudero. 2011. Acoustic properties of Dutch steady-state vowels: contextual effects and a comparison with previous studies. *The 17th International Congress of Phonetic Sciences: Hong Kong, China, August 17-21, 2011: congress proceedings*. 1194-1197. Hong Kong: Department of Chinese, Translation & Linguistics, City University of Hong Kong.

Lew, Robert. 2002. Differences in the scope of obstruent voicing assimilation in learners' English as a consequence of regional variation in Polish. In E. Waniek-Klimczak & P.J. Melia (Eds.), *Accents and speech in teaching English phonetics and phonology*, 243-264. Frankfurt am Main: Lang.

Lobanov, Boris M. 1971. Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America* 49. 606-608.

Locke, J. 1980. The inference of speech perception in the phonologically disordered child. Part II: Some clinically novel procedures, their use, some findings. *Journal of Speech and Hearing Disorders* 45. 445-468.

Marinescu, Irina. 2012. *Native Dialect Effects in Non-native Production and Perception of Vowels.* Toronto, ON: University of Toronto doctoral dissertation.

Mesthrie, Rajend, Joan Swann, Ana Deumert & William L. Leap. 2009. *Introducing Sociolinguistics – Second Edition.* Edinburgh: Edinburgh University Press.

Nooteboom, S.G. 1971. Over de lengte van korte klinkers, lange klinkers en tweeklanken in het Nederlands. *De Nieuwe Taalgids* 64. 396-402.

Nuyts, Jan. 1989. Het Antwerpse vokaalsysteem: Een synchronische en diachronische schets. *Taal en Tongval, Tijdschrift voor Taalvariatie* 41 (1-2). 22-48.

O'brien, Mary Grantham & Laura Catharine Smith. 2010. Role of first language dialect in the production of second language German vowels. *International Review of Applied Linguistics in Language Teaching* 48 (4). 297–330.

Ooms, Miet & Jacques Van Keymeulen. 2005. *Taal in stad en land – Vlaams-Brabants en Antwerps*. Tielt: Lannoo.

Pols, L. C. W., Tromp, H. R. C. & Plomp, R. 1973. Frequency analysis of Dutch vowels from 50 male speakers. *Journal of the Acoustical Society of America* 53 (4), 1093–1101.

Rietveld, Toni, Joop Kerkhoff & Carlos Gussenhoven. 2004. Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics* 32 (3). 349-371.

Taeldeman, Johan. 1985. *De klankstruktuur van het Gentse dialekt: een synchrone beschrijving en een historische en geografische situering*. Wetteren: Universa.

Taeldeman, Johan. 1999. Het Gents. Een eiland in het Oost-Vlaamse dialectgebied. In J. Kruijsen & N. van der Sijs (Eds.), *Structuren in talige variatie in Vlaanderen*. 35-61. Gent: Academie Press.

Thomas, Erik R. & Tyler Kendall. 2007. NORM: The vowel normalization and plotting suite [computer program]. Version 1.1, last modified 19 December 2012. Last consulted 1 August 2013 at http://ncslaap.lib.ncsu.edu/tools/norm/.

Triest, Liesbet. 2009a. Dialectlandschap [website]. http://www.variatielinguistiek.ugent.be/node/43. Last updated 5 November 2009. Last consulted 27 May 2013.

Triest, Liesbet. 2009b. Verwerking van de gegevens: statistiek [website] http://www.variatielinguistiek.ugent.be/node/22. Last updated 9 November 2009. Last consulted 17 July 2013.

Van de Velde, Hans. 2002. Autochtone taalvariatie in het Vlaamse onderwijs: een vergeten zorg? In J. De Caluwe, D. Geeraerts, S. Kroon, V. Mamadouh, R. Soetaert, L. Top & T. Vallen (Eds.), *Taalvariatie en taalbeleid. Bijdragen aan het taalbeleid in Nederland en Vlaanderen*. 131–142. Antwerpen/Apeldoorn: Garant.

Verhoeven, Jo & Christophe Van Bael. 2002a. Acoustic characteristics of monophthong realisation in Southern Standard Dutch. *Antwerp Papers in Linguistics - Phonetic Work in Progress* 100. 149-164.

Verhoeven, Jo & Christophe Van Bael. 2002b. Akoestische kenmerken van de Nederlandse klinkers in drie Vlaamse regio's. *Taal en Tongval, Tijdschrift voor Taalvariatie* 54 (1). 1-23.

Wells, J.C. 1962. *A Study of the Formants of the Pure Vowels of British English*. London: University College London MA dissertation.

Weinreich, U. 1953. *Languages in Contact, Findings and Problems*. The Hague: Mouton.

University of Ghent

Faculty of Arts and Philosophy

2012 - 2013

Appendices

for

Perceiving and producing native and non-native vowels

An experimental study on the effects of first language regional variation

Master thesis submitted in partial fulfillment of the requirements for the degree of
Master of Arts in Linguistics and Literature
English - Latin

Mathijs Debaene

Supervisor: dr. Ellen Simon
Co-supervisor: prof. dr. Mieke Van Herreweghe

Table of contents

Appendix A. Questionnaire results

1 Ghent participants

Informant	Sex	Studies	Date of birth	Age at time of testing	Place of birth	City/town in which you live	City/town in which most childhood/ teenage years were spent
Ghent_1	male	Comparative Modern Literature	10/03/1991	22	Ghent	Zevergem	Ghent
Ghent_2	male	Comparative Modern Literature	14/03/1990	23	Ghent	Ghent	Ghent
Ghent_3	female	Literature and Linguistics: French - Italian	29/10/1993	19	Sint-Agatha-Berchem	Oudenaarde	Ghent
Ghent_4	male	Literature and Linguistics: Greek - Latin	8/04/1993	20	Ghent	Ghent	Ghent
Ghent_5	female	Literature and Linguistics: French - German	27/07/1993	19	Ghent	Ghent	Ghent
Ghent_6	female	Literature and Linguistics: French - German	17/09/1993	19	Ghent	Sint-Amandsberg	Sint-Amandsberg
Ghent_7	female	Literature and Linguistics: Italian - French	6/03/1987	26	Ghent	Ghent	Ghent
Ghent_8	male	Literature and Linguistics: Greek - Latin	17/11/1993	19	Zottegem	Ghent	Ghent
Ghent_9	female	Literature and Linguistics: Greek - Latin	23/07/1993	19	Ghent	Mariakerke	Mariakerke
Ghent_10	female	Literature and Linguistics: Greek-Latin	5/04/1993	20	Ghent	Lochristi	Lochristi
Ghent_11	female	Literature and Linguistics: Greek-Latin	17/07/1992	20	Ghent	Sint-Martens-Latem	Sint-Martens-Latem
Ghent_12	female	Political sciences	29/05/1990	22	Ghent	Ghent	Ghent

1. Ghent participants (continued)

Informant	Other language knowledge	Native language	Language spoken at home	NL of father	NL of mother	Years of English classes at school	Hours of English classes/week	Variety taught in school	Spent time in an English-speaking country?
Ghent_1	French, German	Dutch	Dutch	Dutch	Dutch	5	2	British	no
Ghent_2	French, Spanish	Dutch	Dutch	Dutch	Dutch	6	3	British	Three months in Canterbury (Erasmus)
Ghent_3	French, Italian	Dutch	Dutch/ French	Lingala	Dutch	5	3	British	no
Ghent_4	French, basic Italian	Dutch	Dutch	Dutch	Dutch	5	2	British	no
Ghent_5	French, German, Spanish	Dutch	Dutch/ French	Dutch	Dutch	5	4	British	no
Ghent_6	French, German	Dutch	Dutch	Dutch	Dutch	5	2 to 3	British	no
Ghent_7	French, Italian	Dutch	Dutch	Dutch	Dutch	5	3	British	no
Ghent_8	French	Dutch	Dutch	Dutch	Dutch	5	2	British	no
Ghent_9	French	Dutch	Dutch	Dutch	Dutch	5	2	British	no
Ghent_10	French, basic German	Dutch	Dutch	Dutch	Dutch	5	2 to 3	British	no
Ghent_11	French, basic German, basic Spanish	Dutch	Dutch	Dutch	Dutch	5	2	British	no
Ghent_12	French, Italian, basic Spanish	Dutch	Dutch	Dutch	Dutch	6	2 to 4	British	no

2 Antwerp participants

Informant	Sex	Studies	Date of birth	Age at time of testing	Place of birth	Currently living in...	City/town in which most childhood/teenage years were spent
Antwerp_1	female	Pedagogic Sciences	16/12/1994	18	Mortsel	Brasschaat	Brasschaat
Antwerp_2	female	Literature and Linguistics: Latin - Greek	24/09/1992	20	Antwerp	Antwerp	Antwerp
Antwerp_3	male	Psychology	7/06/1993	19	Antwerp (Wilrijk)	Antwerp (Wilrijk)	Antwerp (Wilrijk)
Antwerp_4	female	Applied Linguistics: French - German	31/10/1994	18	Antwerp (Wilrijk)	Antwerp (Hoboken)	Antwerp (Hoboken)
Antwerp_5	male	Applied Linguistics: Italian - German	26/03/1994	19	Edegem	Mortsel	Mortsel
Antwerp_6	female	Applied Linguistics: French - German	2/08/1994	18	Antwerp (Borgerhout)	Antwerp (Borgerhout)	Antwerp (Borgerhout)
Antwerp_7	female	Legal Practice	25/03/1992	21	Mortsel	Hove	Antwerp
Antwerp_8	female	Speech Therapy	24/11/1993	19	Brasschaat	Schoten	Schoten
Antwerp_9	female	Criminology	12/10/1992	20	Antwerp (Wilrijk)	Mortsel	Mortsel
Antwerp_10	female	Literature and Linguistics: Latin - French	29/03/1995	18	Mortsel	Antwerp (Merksem)	Antwerp (Merksem)
Antwerp_11	female	Psychology	19/05/1992	20	Antwerp (Wilrijk)	Sint-Job-In-'t-Goor	Brasschaat
Antwerp_12	female	Psychology	8/07/1992	20	Brasschaat	Brasschaat	Brasschaat

2. Antwerp participants (continued)

Informant	Other language knowledge	Native language	Language spoken at home	NL of father	NL of mother	Years of English classes at school	Hours of English classes/week	Variety taught in school	Spent time in an English-speaking country?
Antwerp_1	French, Spanish	Dutch	Dutch	French	Dutch	5	2 to 3	British	no
Antwerp_2	French	Dutch	Dutch	Dutch	Dutch	5	2	British	no
Antwerp_3	French, basic German	Dutch	Dutch	Dutch	Dutch	5	3	British	no
Antwerp_4	French, German, basic Spanish	Dutch	Dutch	Dutch	Dutch	5	2 to 3	British	no
Antwerp_5	French, German, Italian	Dutch	Dutch	Dutch	Dutch	5	2	British	no
Antwerp_6	French, German	Dutch	Dutch	Dutch	Dutch	5	2	British	no
Antwerp_7	French	Dutch	Dutch	Dutch	Dutch	6	3	British	no
Antwerp_8	French	Dutch	Dutch	Dutch	Dutch	5	3	British	no
Antwerp_9	French	Dutch	Dutch	Dutch	Dutch	5	3	British	no
Antwerp_10	French, Spanish	Dutch	Dutch	Dutch	Dutch	5	4	British	no
Antwerp_11	French	Dutch	Dutch	Dutch	Dutch	5	3 to 4	British	no
Antwerp_12	French, German	Dutch	Dutch	Dutch	Dutch	6	4	British	no

Appendix B. Perception experiment: Average, standard deviation and median of the target stimuli values

1 Standard Dutch (produced by a native Belgian Standard Dutch speaker)

Standard Dutch			
/h i s/	F1 (Hz)	F2 (Hz)	Duration (ms)
Average	433.4	2662.5	114.7
Standard deviation	8.4	113.6	4.4
Median	434.1	2698.1	113.8
/h ɪ s/	F1 (Hz)	F2 (Hz)	Duration (ms)
Average	446.1	2093.1	78.4
Standard deviation	19.7	86.4	3.8
Median	441.3	2105.9	79.8
/h y s/	F1 (Hz)	F2 (Hz)	Duration (ms)
Average	329.9	2268.9	202.8
Standard deviation	37.4	47.3	18.3
Median	343.6	2251.7	203.8
/h ʏ s/	F1 (Hz)	F2 (Hz)	Duration (ms)
Average	505.4	1711.2	82.2
Standard deviation	10.7	20.0	1.3
Median	507.2	1714.6	82.1

2 RP English (produced by a native RP English speaker)

British (RP) English			
/h iː s/	F1 (Hz)	F2 (Hz)	Duration (ms)
Average	267.8	2618.9	168.0
Standard Deviation	16.5	51.3	5.5
Median	264.8	2601.0	168.3
/h ɪ s/	F1 (Hz)	F2 (Hz)	Duration (ms)
Average	418.5	2235.6	99.8
Standard Deviation	7.8	6.0	12.2
Median	418.6	2236.8	98.1
/h ʊ s/	F1 (Hz)	F2 (Hz)	Duration (ms)
Average	255.4	1132.7	155.6
Standard Deviation	4.8	11.0	5.1
Median	256.7	1132.1	155.1
/h ʌ s/	F1 (Hz)	F2 (Hz)	Duration (ms)
Average	839.8	1377.1	108.4
Standard Deviation	7.5	22.7	4.9
Median	838.7	1382.5	108.3

Appendix C. Perception experiment: individual results

1 Ghent listeners' categorisations

Ghent_Participant_1

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 8**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/aː/	/ɑ/
Hus2 /ʌ/	/aː/	/aː/
Hus3 /ʌ/	/ɑ/	/ɑ/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**5 x /ɑ/, 3 x /aː/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_2

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/ɪ/	/y/
Hies3 /i/	/i/	/i/
Hies4 /i/	/ɪ/	/i/
Total	**/i/ x 5, /ɪ/ x 2, /y/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ɑ/	/aː/
Hus2 /ʌ/	/aː/	/aː/
Hus3 /ʌ/	/aː/	/ɑ/
Hus4 /ʌ/	/ɑ/	/aː/
Total	**5 x /aː/, 3 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_3

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 8**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/øː/	/y/
Huus3 /y/	/øː/	/y/
Huus4 /y/	/øː/	/øː/
Total	**/y/ x 4; /øː/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/ɪ/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**7 x /i/, 1 x /ɪ/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/aː/	/ɑ/
Hus2 /ʌ/	/aː/	/ɑ/
Hus3 /ʌ/	/aː/	/aː/
Hus4 /ʌ/	/aː/	/ɑ/
Total	**5 x /aː/, 3 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_4

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/eː/	/i/
Total	**/i/ x 7; /eː/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ɑ/	/aː/
Hus2 /ʌ/	/aː/	/aː/
Hus3 /ʌ/	/aː/	/ɑ/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**4 x /aː/, 4 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_5

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 8**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ɪ/
Total	**/ʏ/ x 7, /ɪ/ x 1**	
Huus1 /y/	/ʏ/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/ʏ/
Huus4 /y/	/ʏ/	/y/
Total	**/y/ x 5, /ʏ/ x 3**	

English		
Stimulus	1st time	2nd time
Heese1 /i:/	/i/	/i/
Heese2 /i:/	/i/	/i/
Heese3 /i:/	/i/	/i/
Heese4 /i:/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ʏ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**7 x /ɪ/, 1 x /ʏ/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/ɑ/	/ɑ/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**8 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_6

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/ɪ/	/i/
Hies3 /i/	/ɪ/	/i/
Hies4 /i/	/ɪ/	/i/
Total	**/i/ x 5, /ɪ/ x 3**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ʏ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 7, /ʏ/ x 1**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /i:/	/i/	/i/
Heese2 /i:/	/i/	/i/
Heese3 /i:/	/i/	/i/
Heese4 /i:/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/ɑ/	/ɑ/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**8 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_7

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 8**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/aː/	/ɑ/
Hus2 /ʌ/	/aː/	/ɑ/
Hus3 /ʌ/	/aː/	/aː/
Hus4 /ʌ/	/aː/	/ɑ/
Total	**5 x /aː/, 3 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_8

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/ɪ/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 7, /ɪ/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/ɪ/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 7, /ɪ/ x 1**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/aː/	/aː/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/ɑ/	/aː/
Hus4 /ʌ/	/aː/	/ɑ/
Total	**4 x /aː/, 4 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_9

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/eː/	/i/
Total	**/i/ x 7, /eː/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/øː/	/y/
Total	**/y/ x 7, /øː/ x 1**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ʏ/	/ʏ/
Hus2 /ʌ/	/ʏ/	/ʏ/
Hus3 /ʌ/	/ʏ/	/ʏ/
Hus4 /ʌ/	/ʏ/	/ʏ/
Total	**8 x /ʏ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Ghent_Participant_10

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 8**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/aː/	/aː/
Hus2 /ʌ/	/aː/	/aː/
Hus3 /ʌ/	/aː/	/aː/
Hus4 /ʌ/	/aː/	/aː/
Total	**8 x /aː/**	
Hoos1 /ʊ/	/y/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**7 x /u/, 1 x /y/**	

Ghent_Participant_11

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/ɪ/
Hies4 /i/	/i/	/i/
Total	**/i/ x 7, /ɪ/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/i/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**7 x /ɪ/, 1 x /i/**	
Hus1 /ʌ/	/aː/	/aː/
Hus2 /ʌ/	/aː/	/aː/
Hus3 /ʌ/	/aː/	/aː/
Hus4 /ʌ/	/aː/	/aː/
Total	**8 x /aː/**	
Hoos1 /ʊ/	/y/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**7 x /u/, 1 x /y/**	

Ghent_Participant_12

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 8**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/ʏ/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /i:/	/i/	/i/
Heese2 /i:/	/i/	/i/
Heese3 /i:/	/i/	/i/
Heese4 /i:/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ʏ/	/ʏ/
Hus2 /ʌ/	/ʏ/	/ʏ/
Hus3 /ʌ/	/ʏ/	/ʏ/
Hus4 /ʌ/	/ʏ/	/ʏ/
Total	**8 x /ʏ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

2 Antwerp listeners' categorisations

Antwerp_Participant_1

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/ɪ/	/i/
Hies2 /i/	/i/	/ɪ/
Hies3 /i/	/ɪ/	/ɪ/
Hies4 /i/	/ɪ/	/i/
Total	**/ɪ/ x 5, /i/ x 3**	
His1 /ɪ/	/i/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/i/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 6, /i/ x 2**	
Hus1 /ʏ/	/øː/	/ʏ/
Hus2 /ʏ/	/ʏ/	/øː/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/øː/	/øː/
Total	**/øː/ x 4, /ʏ/ x 4**	
Huus1 /y/	/ʏ/	/y/
Huus2 /y/	/y/	/ʏ/
Huus3 /y/	/y/	/ʏ/
Huus4 /y/	/y/	/y/
Total	**/y/ x 5, /ʏ/ x 3**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/ɪ/	/i/
Total	**7 x /i/, 1 x /ɪ/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/ɑ/	/aː/
Hus3 /ʌ/	/ɑ/	/aː/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**6 x /ɑ/, 2 x /aː/**	
Hoos1 /ʊ/	/ʏ/	/u/
Hoos2 /ʊ/	/u/	/ʏ/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/ʏ/
Total	**5 x /u/, 3 x /ʏ/**	

Antwerp_Participant_2

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/ɪ/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 7, /ɪ/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /i:/	/i/	/i/
Heese2 /i:/	/i/	/i/
Heese3 /i:/	/i/	/i/
Heese4 /i:/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/ɑ/	/ɑ/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**8 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Antwerp_Participant_3

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/ɪ/	/ɪ/
Hies2 /i/	/ɪ/	/ɪ/
Hies3 /i/	/ɪ/	/ɪ/
Hies4 /i/	/ɪ/	/i/
Total	**/ɪ/ x 7, /i/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/ʏ/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 7, /ʏ/ x 1**	

English		
Stimulus	1st time	2nd time
Heese1 /i:/	/i/	/i/
Heese2 /i:/	/ɪ/	/i/
Heese3 /i:/	/i/	/i/
Heese4 /i:/	/ɪ/	/i/
Total	**6 x /i/, 2 x /ɪ/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/ɑ/	/ɑ/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**8 x /ɑ/**	
Hoos1 /ʊ/	/u/	/ʏ/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**7 x /u/, 1 x /ʏ/**	

Antwerp_Participant_4

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/i/
Hies3 /i/	/i/	/ɪ/
Hies4 /i/	/i/	/ɪ/
Total	**/i/ x 6, /ɪ/ x 2**	
His1 /ɪ/	/i/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 7, /i/ x 1**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/øː/	/y/
Huus4 /y/	/øː/	/y/
Total	**/y/ x 6, /øː/ x 2**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/e/	/ɪ/
His3 /ɪ/	/e/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**6 x /ɪ/, 2 x /e/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/aː/	/ɑ/
Hus3 /ʌ/	/aː/	/aː/
Hus4 /ʌ/	/aː/	/ɑ/
Total	**4 x /aː/, 4 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/ʏ/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/ʏ/
Total	**6 x /u/, 2 x /ʏ/**	

Antwerp_Participant_5

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/ɪ/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 7, /ɪ/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /i:/	/i/	/i/
Heese2 /i:/	/i/	/ɪ/
Heese3 /i:/	/i/	/i/
Heese4 /i:/	/i/	/i/
Total	**7 x /i/, 1 x /ɪ/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ʏ/	/ʏ/
Hus2 /ʌ/	/ʏ/	/ʏ/
Hus3 /ʌ/	/ʏ/	/ʏ/
Hus4 /ʌ/	/ʏ/	/ʏ/
Total	**8 x /ʏ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Antwerp_Participant_6

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/i/	/ɪ/
Hies3 /i/	/i/	/i/
Hies4 /i/	/i/	/i/
Total	**/i/ x 7, /ɪ/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/i/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 7, /i/ x 1**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /i/**	
Hus1 /ʌ/	/aː/	/ɑ/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/aː/	/ɑ/
Hus4 /ʌ/	/aː/	/ɑ/
Total	**5 x /ɑ/, 3 x /aː/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/ʏ/	/u/
Total	**7 x /u/, 1 x /ʏ/**	

Antwerp_Participant_7

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/ɪ/
Hies2 /i/	/ɪ/	/ɪ/
Hies3 /i/	/ɪ/	/ɪ/
Hies4 /i/	/i/	/ɪ/
Total	**/ɪ/ x 6, /i/ x 2**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/øː/	/ʏ/
Total	**/ʏ/ x 7, /øː/ x 1**	
Huus1 /y/	/øː/	/øː/
Huus2 /y/	/øː/	/øː/
Huus3 /y/	/øː/	/øː/
Huus4 /y/	/øː/	/øː/
Total	**/øː/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/ɪ/
Heese3 /iː/	/ɪ/	/i/
Heese4 /iː/	/i/	/i/
Total	**6 x /i/, 2 x /ɪ/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/ɑ/	/ɑ/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**8 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/øː/	/u/
Hoos3 /ʊ/	/u/	/ʏ/
Hoos4 /ʊ/	/u/	/u/
Total	**6 x /u/, 1 x /øː/, 1 x /ʏ/**	

Antwerp_Participant_8

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/ɪ/
Hies2 /i/	/ɪ/	/i/
Hies3 /i/	/i/	/i/
Hies4 /i/	/ɪ/	/i/
Total	**/i/ x 5, /ɪ/ x 3**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/ɪ/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 7, /ɪ/ x 1**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/aː/	/aː/
Hus2 /ʌ/	/aː/	/aː/
Hus3 /ʌ/	/aː/	/aː/
Hus4 /ʌ/	/aː/	/aː/
Total	**8 x /aː/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**8 x /u/**	

Antwerp_Participant_9

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/ɪ/	/i/
Hies2 /i/	/i/	/ɪ/
Hies3 /i/	/i/	/ɪ/
Hies4 /i/	/i/	/ɪ/
Total	**/i/ x 4, /ɪ/ x 4**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/i/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 7, /i/ x 1**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/ʏ/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 7, /ʏ/ x 1**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/i/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**7 x /ɪ/, 1 x /i/**	
Hus1 /ʌ/	/aː/	/aː/
Hus2 /ʌ/	/aː/	/aː/
Hus3 /ʌ/	/aː/	/aː/
Hus4 /ʌ/	/aː/	/aː/
Total	**8 x /aː/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/ʏ/
Total	**7 x /u/, 1 x /ʏ/**	

Antwerp_Participant_10

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/ɪ/
Hies2 /i/	/ɪ/	/i/
Hies3 /i/	/ɪ/	/i/
Hies4 /i/	/ɪ/	/ɪ/
Total	**/ɪ/ x 5, /i/ x 3**	
His1 /ɪ/	/ɪ/	/i/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 7, /i/ x 1**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/øː/	/i/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 6, /øː/ x 1, /i/ x 1**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/i/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/i/	/i/
Heese4 /iː/	/i/	/i/
Total	**8 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/ɑ/	/ɑ/
Hus4 /ʌ/	/aː/	/ɑ/
Total	**7 x /ɑ/, 1 x /aː/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/ʏ/	/u/
Hoos3 /ʊ/	/ʏ/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**6 x /u/, 2 x /ʏ/**	

Antwerp_Participant_11

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/ɪ/	/ɪ/
Hies2 /i/	/ɪ/	/ɪ/
Hies3 /i/	/ɪ/	/ɪ/
Hies4 /i/	/eː/	/i/
Total	**/ɪ/ x 6, /i/ x 1, /eː/ x 1**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/e/	/e/
His3 /ɪ/	/e/	/ɪ/
His4 /ɪ/	/e/	/ɪ/
Total	**/e/ x 4, /ɪ/ x 4**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/ʏ/	/y/
Huus2 /y/	/y/	/øː/
Huus3 /y/	/y/	/y/
Huus4 /y/	/øː/	/øː/
Total	**/y/ x 4, /øː/ x 3, /ʏ/ x 1**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/ɪ/	/ɪ/
Heese2 /iː/	/i/	/i/
Heese3 /iː/	/ɪ/	/ɪ/
Heese4 /iː/	/ɪ/	/i/
Total	**5 x /ɪ/, 3 x /i/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/e/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**7 x /ɪ/, 1 x /e/**	
Hus1 /ʌ/	/ɑ/	/ɑ/
Hus2 /ʌ/	/ɑ/	/ɑ/
Hus3 /ʌ/	/ɑ/	/ɑ/
Hus4 /ʌ/	/ɑ/	/ɑ/
Total	**8 x /ɑ/**	
Hoos1 /ʊ/	/u/	/u/
Hoos2 /ʊ/	/ʏ/	/u/
Hoos3 /ʊ/	/u/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**7 x /u/, 1 x/ʏ/**	

Antwerp_Participant_12

Dutch		
Stimulus	1st time	2nd time
Hies1 /i/	/i/	/i/
Hies2 /i/	/ɪ/	/i/
Hies3 /i/	/ɪ/	/i/
Hies4 /i/	/ɪ/	/i/
Total	**/i/ x 5, /ɪ/ x 3**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**/ɪ/ x 8**	
Hus1 /ʏ/	/ʏ/	/ʏ/
Hus2 /ʏ/	/ʏ/	/ʏ/
Hus3 /ʏ/	/ʏ/	/ʏ/
Hus4 /ʏ/	/ʏ/	/ʏ/
Total	**/ʏ/ x 8**	
Huus1 /y/	/y/	/y/
Huus2 /y/	/y/	/y/
Huus3 /y/	/y/	/y/
Huus4 /y/	/y/	/y/
Total	**/y/ x 8**	

English		
Stimulus	1st time	2nd time
Heese1 /iː/	/i/	/ɪ/
Heese2 /iː/	/i/	/ɪ/
Heese3 /iː/	/eː/	/i/
Heese4 /iː/	/i/	/i/
Total	**5 x /i/, 2 x /ɪ/, 1 x /eː/**	
His1 /ɪ/	/ɪ/	/ɪ/
His2 /ɪ/	/ɪ/	/ɪ/
His3 /ɪ/	/ɪ/	/ɪ/
His4 /ɪ/	/ɪ/	/ɪ/
Total	**8 x /ɪ/**	
Hus1 /ʌ/	/aː/	/aː/
Hus2 /ʌ/	/ɑ/	/aː/
Hus3 /ʌ/	/aː/	/aː/
Hus4 /ʌ/	/aː/	/aː/
Total	**7 x /aː/, 1 x /ɑ/**	
Hoos1 /ʊ/	/u/	/ʏ/
Hoos2 /ʊ/	/u/	/u/
Hoos3 /ʊ/	/ʏ/	/u/
Hoos4 /ʊ/	/u/	/u/
Total	**6 x /u/, 2 x /ʏ/**	

Appendix D. Perception experiment: SPSS outputs

1 Dutch-to-Dutch perception task

hiesDutch * regio Crosstabulation

			regio		
			1,00	2,00	Total
hiesDutch	1,00	Count	51	86	137
		Expected Count	68,5	68,5	137,0
		% within hiesDutch	37,2%	62,8%	100,0%
	2,00	Count	44	7	51
		Expected Count	25,5	25,5	51,0
		% within hiesDutch	86,3%	13,7%	100,0%
	3,00	Count	1	2	3
		Expected Count	1,5	1,5	3,0
		% within hiesDutch	33,3%	66,7%	100,0%
	11,00	Count	0	1	1
		Expected Count	,5	,5	1,0
		% within hiesDutch	0,0%	100,0%	100,0%
Total		Count	96	96	192
		Expected Count	96,0	96,0	192,0
		% within hiesDutch	50,0%	50,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	37,118[a]	3	,000	,000		
Likelihood Ratio	40,674	3	,000	,000		
Fisher's Exact Test	39,505			,000		
Linear-by-Linear Association	4,453[b]	1	,035	,017	,009	,005
N of Valid Cases	192					

a. 4 cells (50,0%) have expected count less than 5. The minimum expected count is ,50.

b. The standardized statistic is -2,110.

SPSS output for the statistical analysis of Dutch /his/ stimuli categorisation in the Dutch perception task. The appropriate statistical significance test, its value, degrees of freedom (only relevant for the Pearson chi-squared test) and p-value (i.c. 'Exact Sig. (2-sided)) are indicated with a green box.

hisDutch * regio Crosstabulation

			regio		
			1,00	2,00	Total
hisDutch	1,00	Count	6	0	6
		Expected Count	3,0	3,0	6,0
		% within hisDutch	100,0%	0,0%	100,0%
	2,00	Count	86	95	181
		Expected Count	90,5	90,5	181,0
		% within hisDutch	47,5%	52,5%	100,0%
	4,00	Count	4	0	4
		Expected Count	2,0	2,0	4,0
		% within hisDutch	100,0%	0,0%	100,0%
	12,00	Count	0	1	1
		Expected Count	,5	,5	1,0
		% within hisDutch	0,0%	100,0%	100,0%
Total		Count	96	96	192
		Expected Count	96,0	96,0	192,0
		% within hisDutch	50,0%	50,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	11,448[a]	3	,010	,002		
Likelihood Ratio	15,697	3	,001	,002		
Fisher's Exact Test	11,162			,002		
Linear-by-Linear Association	,525[b]	1	,469	,732	,366	,076
N of Valid Cases	192					

a. 6 cells (75,0%) have expected count less than 5. The minimum expected count is ,50.

b. The standardized statistic is ,725.

SPSS output for the statistical analysis of Dutch /hɪs/ stimuli categorisation in the Dutch perception task. The appropriate statistical significance test, its value, degrees of freedom (only relevant for the Pearson chi-squared test) and p-value (i.c. 'Exact Sig. (2-sided)) are indicated with a green box.

huusDutch * regio Crosstabulation

			regio		
			1,00	2,00	Total
huusDutch	1,00	Count	1	0	1
		Expected Count	,5	,5	1,0
		% within huusDutch	100,0%	0,0%	100,0%
	2,00	Count	1	1	2
		Expected Count	1,0	1,0	2,0
		% within huusDutch	50,0%	50,0%	100,0%
	9,00	Count	14	5	19
		Expected Count	9,5	9,5	19,0
		% within huusDutch	73,7%	26,3%	100,0%
	11,00	Count	74	87	161
		Expected Count	80,5	80,5	161,0
		% within huusDutch	46,0%	54,0%	100,0%
	12,00	Count	6	3	9
		Expected Count	4,5	4,5	9,0
		% within huusDutch	66,7%	33,3%	100,0%
Total		Count	96	96	192
		Expected Count	96,0	96,0	192,0
		% within huusDutch	50,0%	50,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	7,313[a]	4	,120	,077		
Likelihood Ratio	7,895	4	,095	,103		
Fisher's Exact Test	7,409			,066		
Linear-by-Linear Association	1,884[b]	1	,170	,199	,099	,018
N of Valid Cases	192					

a. 6 cells (60,0%) have expected count less than 5. The minimum expected count is ,50.

b. The standardized statistic is 1,372.

SPSS output for the statistical analysis of Dutch /hys/ stimuli categorisation in the Dutch perception task. The appropriate statistical significance test, its value, degrees of freedom (only relevant for the Pearson chi-squared test) and p-value (i.c. 'Exact Sig. (2-sided)) are indicated with a green box.

husDutch * regio Crosstabulation

			regio		
			1,00	2,00	Total
husDutch	2,00	Count	0	1	1
		Expected Count	,5	,5	1,0
		% within husDutch	0,0%	100,0%	100,0%
	9,00	Count	5	0	5
		Expected Count	2,5	2,5	5,0
		% within husDutch	100,0%	0,0%	100,0%
	12,00	Count	91	95	186
		Expected Count	93,0	93,0	186,0
		% within husDutch	48,9%	51,1%	100,0%
Total		Count	96	96	192
		Expected Count	96,0	96,0	192,0
		% within husDutch	50,0%	50,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6,086[a]	2	,048	,059		
Likelihood Ratio	8,404	2	,015	,059		
Fisher's Exact Test	5,891			,059		
Linear-by-Linear Association	,175[b]	1	,675	,843	,421	,015
N of Valid Cases	192					

a. 4 cells (66,7%) have expected count less than 5. The minimum expected count is ,50.

b. The standardized statistic is ,419.

SPSS output for the statistical analysis of Dutch /hʏs/ stimuli categorisation in the Dutch perception task. The appropriate statistical significance test, its value, degrees of freedom (only relevant for the Pearson chi-squared test) and p-value (i.c. 'Exact Sig. (2-sided)) are indicated with a green box.

2 English-to-Dutch perception task

heeseEnglish * regio Crosstabulation

			regio		
			1,00	2,00	Total
heeseEnglish	1,00	Count	82	95	177
		Expected Count	88,5	88,5	177,0
		% within heeseEnglish	46,3%	53,7%	100,0%
	2,00	Count	13	1	14
		Expected Count	7,0	7,0	14,0
		% within heeseEnglish	92,9%	7,1%	100,0%
	3,00	Count	1	0	1
		Expected Count	,5	,5	1,0
		% within heeseEnglish	100,0%	0,0%	100,0%
Total		Count	96	96	192
		Expected Count	96,0	96,0	192,0
		% within heeseEnglish	50,0%	50,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	12,241[a]	2	,002	,001		
Likelihood Ratio	14,545	2	,001	,001		
Fisher's Exact Test	12,925			,001		
Linear-by-Linear Association	11,699[b]	1	,001	,001	,000	,000
N of Valid Cases	192					

a. 2 cells (33,3%) have expected count less than 5. The minimum expected count is ,50.

b. The standardized statistic is -3,420.

SPSS output for the statistical analysis of English /hi:s/ stimuli categorisation in the cross-language perception task. The appropriate statistical significance test, its value, degrees of freedom (only relevant for the Pearson chi-squared test) and p-value (i.c. 'Exact Sig. (2-sided)) are indicated with a green box.

hisEnglish * regio Crosstabulation

			regio		Total
			1,00	2,00	
hisEnglish	1,00	Count	1	1	2
		Expected Count	1,0	1,0	2,0
		% within hisEnglish	50,0%	50,0%	100,0%
	2,00	Count	92	94	186
		Expected Count	93,0	93,0	186,0
		% within hisEnglish	49,5%	50,5%	100,0%
	4,00	Count	3	0	3
		Expected Count	1,5	1,5	3,0
		% within hisEnglish	100,0%	0,0%	100,0%
	12,00	Count	0	1	1
		Expected Count	,5	,5	1,0
		% within hisEnglish	0,0%	100,0%	100,0%
Total		Count	96	96	192
		Expected Count	96,0	96,0	192,0
		% within hisEnglish	50,0%	50,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	4,022[a]	3	,259	,246		
Likelihood Ratio	5,567	3	,135	,246		
Fisher's Exact Test	3,806			,246		
Linear-by-Linear Association	,141[b]	1	,707	,968	,484	,031
N of Valid Cases	192					

a. 6 cells (75,0%) have expected count less than 5. The minimum expected count is ,50.

b. The standardized statistic is ,375.

SPSS output for the statistical analysis of English /hɪs/ stimuli categorisation in the cross-language perception task. The appropriate statistical significance test, its value, degrees of freedom (only relevant for the Pearson chi-squared test) and p-value (i.c. 'Exact Sig. (2-sided)) are indicated with a green box.

hoosEnglish * regio Crosstabulation

			regio		Total
			1,00	2,00	
hoosEnglish	9,00	Count	1	0	1
		Expected Count	,5	,5	1,0
		% within hoosEnglish	100,0%	0,0%	100,0%
	10,00	Count	81	94	175
		Expected Count	87,5	87,5	175,0
		% within hoosEnglish	46,3%	53,7%	100,0%
	11,00	Count	0	2	2
		Expected Count	1,0	1,0	2,0
		% within hoosEnglish	0,0%	100,0%	100,0%
	12,00	Count	14	0	14
		Expected Count	7,0	7,0	14,0
		% within hoos English	100,0%	0,0%	100,0%
Total		Count	96	96	192
		Expected Count	96,0	96,0	192,0
		% within hoosEnglish	50,0%	50,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	17,966[a]	3	,000	,000		
Likelihood Ratio	24,534	3	,000	,000		
Fisher's Exact Test	19,841			,000		
Linear-by-Linear Association	11,383[b]	1	,001	,001	,000	,000
N of Valid Cases	192					

a. 4 cells (50,0%) have expected count less than 5. The minimum expected count is ,50.

b. The standardized statistic is -3,374.

SPSS output for the statistical analysis of English /hʊs/ stimuli categorisation in the cross-language perception task. The appropriate statistical significance test, its value, degrees of freedom (only relevant for the Pearson chi-squared test) and p-value (i.c. 'Exact Sig. (2-sided)) are indicated with a green box.

husEnglish * regio Crosstabulation

			regio		
			1,00	2,00	Total
husEnglish	5,00	Count	33	42	75
		Expected Count	37,5	37,5	75,0
		% within husEnglish	44,0%	56,0%	100,0%
	6,00	Count	55	38	93
		Expected Count	46,5	46,5	93,0
		% within husEnglish	59,1%	40,9%	100,0%
	12,00	Count	8	16	24
		Expected Count	12,0	12,0	24,0
		% within husEnglish	33,3%	66,7%	100,0%
Total		Count	96	96	192
		Expected Count	96,0	96,0	192,0
		% within husEnglish	50,0%	50,0%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6,854[a]	2	,032	,036		
Likelihood Ratio	6,926	2	,031	,036		
Fisher's Exact Test	6,794			,036		
Linear-by-Linear Association	1,655[b]	1	,198	,212	,106	,012
N of Valid Cases	192					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 12,00.

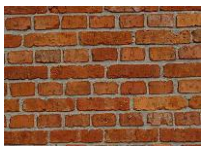
b. The standardized statistic is 1,286.

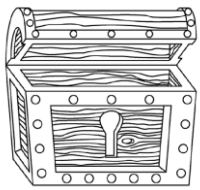
SPSS output for the statistical analysis of English /hʌs/ stimuli categorisation in the cross-language perception task. The appropriate statistical significance test, its value, degrees of freedom (only relevant for the Pearson chi-squared test) and p-value (i.c. 'Asymp.Sig. (2-sided)) are indicated with a green box.

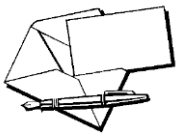
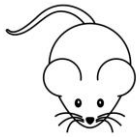
Appendix E. Production experiment: Visual Stimuli and Carrier sentences

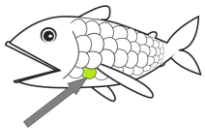
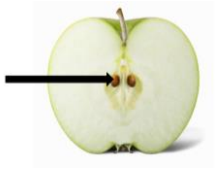
1 Dutch production task (in order or appearance; distracters are indicated in cursive)

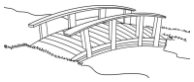
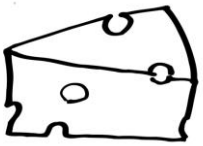
schip /sxɪp/ 'ship'	vuur /vyr/'fire'	*stoel /stul/, 'chair'*	kus /kʏs/ 'kiss'
			
Hij is matroos op een... ('He is a sailor on a ...')	Ze verbrandde zich aan het... ('She burnt herself on the...')	*Een... heeft een leuning ('A... has a back.')*	Ze gaf de jongen een ... ('She gave the boy a...')

dief /dif/ 'thief'	muur /myr/ 'wall'	*geel /ɣe:l/, 'yellow'*	*beer /be:r/, 'bear'*
			
Ze werden bestolen door een... ('They were robbed by a...')	Ze klom over een... ('She climbed a...')	*Citroenen zijn... ('Lemons are...')*	*Hij werd aangevallen door een... ('He was attacked by a...')*

kist /kɪst/ 'chest'	*glas /ɣlas/, 'glass'*	*huis /hœys/ 'house'*	put /pʏt/ 'pit'
			
Hij vond een schat in de... ('He found a treasure in the...')	*Het... viel op de grond en brak. ('The ... fell on the ground and broke.')*	*Ze woont in een groot... ('She lives in a big...')*	Ze vielen in een... ('They fell in a...')

brief /brif/‘letter’	stuur /styr/ ‘steering wheel’	*bal /bal/, ‘ball’*	*muis /mœys/, ‘mouse’*
			
Hij schreef haar een... (‘He wrote her...’)	Hij verloor de controle over het ... (‘He lost control of the...’)	*De... rolde in het doel. (‘The... rolled in the goal.’)*	*Het gepiep kwam van een... (‘The peeping noise came from a...’)*

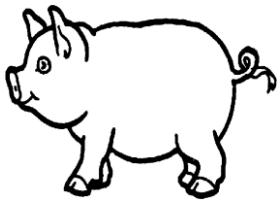
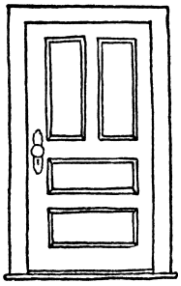
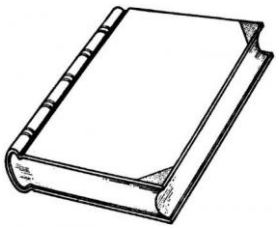
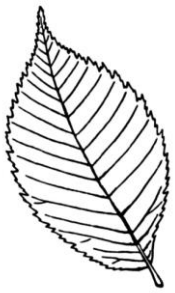
schub /sxʏp/ ‘scale’	duur /dyr/‘expensive’	*aap /a:p/ ‘monkey’*	pit /pɪt/‘kernel’
			
De vis miste een... (‘The fish missed a...’)	Die jas was... (‘That coat was...’)	*De... at een banaan. (‘The... ate a banana.’)*	In de appel zit een... (‘The apple contains a a...’)

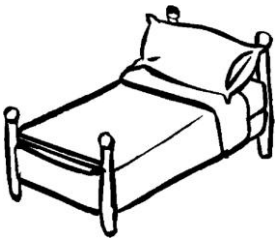
kies /kis/ ‘molar’	brug /brʏx/‘bridge’	sik /sɪk/ ‘goatee’	*kaas /ka:s/, ‘cheese’*
			
Er zat een gaatje in haar... (‘There was a hole in her...’)	Hij wandelde over de... (‘He walked across...’)	De geit heeft een... (‘The goat has a...’)	*De... stond al op tafel. (‘The... was already on the table.’)*

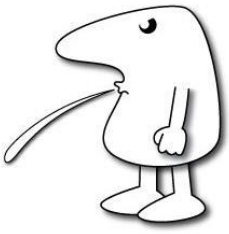
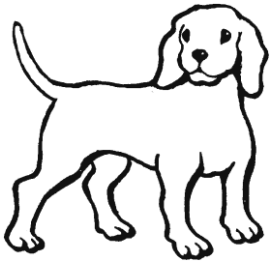
ziek /zik/ ‘ill’

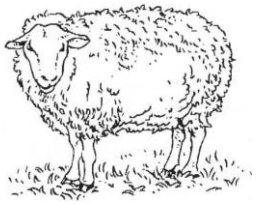
Deze man is... (‘This man is...’)

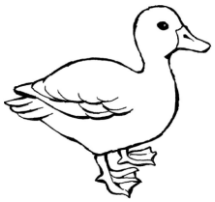
2 English production task (in order or appearance; distracters are indicated in cursive)

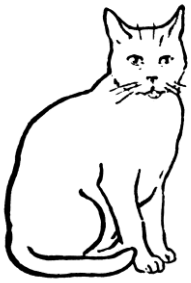
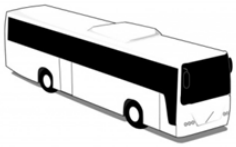
pig /p ɪ g/	*door /d ɔ:/*	book /b ʊ k/	leaf /l i: f/
			
He had a little farm with one...	*She locked the...*	The girl was reading a...	He stepped on a crunchy ...

hug /h ʌ g/	key /k i:/	*bed /b ɛ d/*	cook /k ʊ k/
			
She gave him a...	He turned the ...	The... was empty.	He's a very skilled...

spit /s p ɪ t/	*dog /d ɒ g/*	*shoe /ʃ u:/*	cup /k ʌ p/
			
There was... on the floor.	The... barked at him.	He tied the laces of his...	He drank tea from a ...

grass /g r ɑ: s/	foot /f ʊ t/	sheep /ʃ i: p/	*bat /b æ t/*
			
The... is green.	There was a splinter in his left...	Wool comes from a...	He was scared by a...

fist /f ɪ s t/	*rose /r ə ʊ z/*	duck /d ʌ k/	peace /p i: s/
			
She punched him with her ...	A ... is a type of flower.	He fed bread crumbs to a ...	The Nobel Prize for...

six /s ɪ k s/	*red /r ɛ d/*	*cat /k æ t/*	bus /b ʌ s/
			
He had to be home by...	The wine was a dark...	He gave food to the ...	She always takes the...

hook /h ʊ k/

There was a fish on his...

Appendix F. Production experiment: individual values

1 Ghent participants

Ghent_Participant_1

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	387,8	1798,6	56,5
Sik	424,6	1994,1	59,3
Pit	439,4	1679,1	57,8
Kist	437,3	1815,2	77,4
Average	422,275	1821,75	62,75
Standard dev.	23,89524011	129,9132659	9,833446327
Median	430,95	1806,9	58,55
Kies	255,5	2105,6	78,5
Brief	321,6	2175,7	77,9
Dief	234,3	1933,2	74,2
Ziek	302,7	2126,1	62,1
Average	278,525	2085,15	73,175
Standard dev.	40,52113646	105,4880878	7,624248597
Median	279,1	2115,85	76,05
Put	401,3	1456,1	75,1
Schub	404,8	1559,7	81,4
Kus	413,6	1631,1	76,2
Brug	420,8	1478,5	81,9
Average	410,125	1531,35	78,65
Standard dev.	8,799005626	80,01972673	3,499047489
Median	409,2	1519,1	78,8
Muur	265,3	1803,2	145,9
Stuur	276,6	1873,1	162,8
Vuur	232,8	1871,6	159,3
Duur	214,9	1879,8	164,1
Average	247,4	1856,925	158,025
Standard dev.	28,53337227	35,99364527	8,333616662
Median	249,05	1872,35	161,05

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	407,8	1743,2	87,8
Fist	449,1	1732,6	86,9
Six	417,3	2052,1	75,3
Pig	407,8	2045,2	145,4
Average	420,5	1893,275	98,85
Standard dev.	19,58553888	179,4858838	31,55106971
Median	412,55	1894,2	87,35
Key	278,2	2291,6	164,8
Sheep	352,4	2365	146,3
Peace	285,4	2062	157,8
Leaf	307	2306,6	179,1
Average	305,75	2256,3	162
Standard dev.	33,42110112	133,347116	13,71592748
Median	296,2	2299,1	161,3
Hug	556,8	1171,9	151,3
Duck	594,6	1241,4	113,6
Cup	603,1	1141,7	99,7
Bus	581,2	1059,6	124,5
Average	583,925	1153,65	122,275
Standard dev.	20,20583662	75,32447588	21,85015255
Median	587,9	1156,8	119,05
Book	386	1139,8	82,9
Foot	353,6	1194,7	85,2
Hook	338,4	1207,1	96,2
Cook	353,1	1005,7	75,3
Average	357,775	1136,825	84,9
Standard dev.	20,09417411	92,17864449	8,639830245
Median	353,35	1167,25	84,05

Ghent_Participant_2

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	426,4	1676,2	61,3
Sik	368,5	2007,2	64,1
Pit	398,2	1677,5	63,6
Kist	422,1	1726,9	58,9
Average	403,8	1771,95	61,975
Standard dev.	26,60263145	158,5990017	2,385197406
Median	410,15	1702,2	62,45
Kies	252,8	2341,1	61,4
Brief	283,1	2074,8	70,4
Dief	263,5	2057,2	62,1
Ziek	283,7	2394,4	66,4
Average	270,775	2216,875	65,075
Standard dev.	15,2204632	175,716047	4,182005101
Median	273,3	2207,95	64,25
Put	402,7	1417,7	68
Schub	442,3	1649,5	63,8
Kus	396,8	1570,5	73,1
Brug	389,9	1582,9	76,7
Average	407,925	1555,15	70,4
Standard dev.	23,50608077	97,97977682	5,66568619
Median	399,75	1576,7	70,55
Muur	273,2	1796,3	180,8
Stuur	265,1	2122,7	164,8
Vuur	237,6	2089,7	157,8
Duur	242,5	1988,2	176,2
Average	254,6	1999,225	169,9
Standard dev.	17,23968291	146,8921458	10,50269807
Median	253,8	2038,95	170,5

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	404,9	1682,9	62,5
Fist	419,4	1718,1	62,6
Six	367,4	2080,5	51
Pig	359	2118,1	113,5
Average	387,675	1899,9	72,4
Standard dev.	29,07752569	231,2054209	27,93575964
Median	386,15	1899,3	62,55
Key	244,9	2348,5	133,5
Sheep	279,7	2193,6	139,2
Peace	268,3	2517,1	162
Leaf	264,5	2283	148,1
Average	264,35	2335,55	145,7
Standard dev.	14,48620033	136,674077	12,41692393
Median	266,4	2315,75	143,65
Hug	565,2	1278,1	140,1
Duck	579	1255,3	88,9
Cup	636,9	1219,7	69,1
Bus	463,3	1054,1	115
Average	561,1	1201,8	103,275
Standard dev.	72,22118803	101,3568613	30,91993693
Median	572,1	1237,5	101,95
Book	374,5	981,5	59,6
Foot	352,1	1602,1	60,4
Hook	363,4	998,8	64,6
Cook	361,8	997,5	65,6
Average	362,95	1144,975	62,55
Standard dev.	9,176963913	304,8515199	2,990540642
Median	362,6	998,15	62,5

Ghent_Participant_3

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	480,2	2164,2	45,6
Sik	420,5	2329,7	54,1
Pit	477,8	1990,1	55,3
Kist	478	2155,2	64,8
Average	464,125	2159,8	54,95
Standard dev.	29,10365097	138,6898458	7,858965157
Median	477,9	2159,7	54,7
Kies	300,7	2620,8	64,9
Brief	298,3	2683,7	93,4
Dief	267,7	2785,7	77,7
Ziek	311,9	2850,3	83,8
Average	294,65	2735,125	79,95
Standard dev.	18,91903803	102,5310481	11,93440405
Median	299,5	2734,7	80,75
Put	423,7	1674,4	75,1
Schub	434,5	1662,2	56,3
Kus	440	1799,9	75
Brug	442,5	1697,3	62,6
Average	435,175	1708,45	67,25
Standard dev.	8,348003753	62,67878429	9,36678529
Median	437,25	1685,85	68,8
Muur	315,3	1709	162,6
Stuur	322,4	2077,5	162,8
Vuur	343,1	2243,3	154,3
Duur	309,4	2035,4	145,7
Average	322,55	2016,3	156,35
Standard dev.	14,69478365	223,6608593	8,12998565
Median	318,85	2056,45	158,45

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	479,3	2011,5	78,7
Fist	445,4	2081,7	74,2
Six	394,5	2353,1	85,9
Pig	465,4	2515,9	97,3
Average	446,15	2240,55	84,025
Standard dev.	37,13834496	235,3595476	10,07682986
Median	455,4	2217,4	82,3
Key	313,8	3006,3	151,4
Sheep	359,1	2908,6	137,7
Peace	334,7	2490,6	149,4
Leaf	319,3	2840,3	121,3
Average	331,725	2811,45	139,95
Standard dev.	20,28075853	224,4858496	13,8237718
Median	327	2874,45	143,55
Hug	747,4	1272	130
Duck	643,4	1505,6	85,6
Cup	678,9	1330,6	91,2
Bus	721,9	1336,1	114,5
Average	697,9	1361,075	105,325
Standard dev.	46,03440742	100,621912	20,66903884
Median	700,4	1333,35	102,85
Book	453,8	1021,2	87,3
Foot	493,8	1263,8	70,3
Hook	435,3	1066,5	70,2
Cook	411,4	937	73,9
Average	448,575	1072,125	75,425
Standard dev.	34,78892305	138,5918077	8,101594493
Median	444,55	1043,85	72,1

Ghent_Participant_4

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	472,5	1735,4	82,6
Sik	449,4	1974,4	70,6
Pit	419,6	1779,8	62,7
Kist	463,9	1748,5	91,1
Average	451,35	1809,525	76,75
Standard dev.	23,21414798	111,4837918	12,58795721
Median	456,65	1764,15	76,6
Kies	219,6	2040,2	88,6
Brief	314,4	2339,8	88,7
Dief	267,9	2112,7	78,9
Ziek	309,2	2168,6	77,7
Average	277,775	2165,325	83,475
Standard dev.	44,01048171	127,6426098	5,995762392
Median	288,55	2140,65	83,75
Put	495,3	1399,5	65
Schub	541,7	1485,2	76,2
Kus	542,7	1472,8	97,3
Brug	421,7	1556,9	103,9
Average	500,35	1478,6	85,6
Standard dev.	56,90538932	64,45903092	18,11537101
Median	518,5	1479	86,75
Muur	297,2	1918,1	186,3
Stuur	284,9	2039,2	183,4
Vuur	283,7	1864,4	210,8
Duur	274,4	1898,2	200,7
Average	285,05	1929,975	195,3
Standard dev.	9,36108968	76,11580541	12,80650876
Median	284,3	1908,15	193,5

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	440,8	2003,1	60,3
Fist	377,4	2147,9	74,6
Six	405	2215,2	87,6
Pig	365,5	2168,7	104,5
Average	397,175	2133,725	81,75
Standard dev.	33,46006326	91,51547683	18,82383241
Median	391,2	2158,3	81,1
Key	230,1	2252,1	199,1
Sheep	255,5	2274,9	131,9
Peace	253,1	2291,9	138,1
Leaf	303,8	2232,6	119,6
Average	260,625	2262,875	147,175
Standard dev.	30,97712866	25,94691697	35,46015369
Median	254,3	2263,5	135
Hug	566,4	1309,4	92,7
Duck	669,4	1254,5	97,6
Cup	727,3	1202,5	91,3
Bus	693,6	1190,1	84,1
Average	664,175	1239,125	91,425
Standard dev.	69,37301949	54,52769174	5,580546568
Median	681,5	1228,5	92
Book	449,3	771,7	68,7
Foot	400,3	901,9	73,5
Hook	465,6	792	76,4
Cook	374,5	860,5	92
Average	422,425	831,525	77,65
Standard dev.	42,31889846	60,36954944	10,07984788
Median	424,8	826,25	74,95

Ghent_Participant_5

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	477,3	2043,9	73,4
Sik	415,4	2287,4	76,3
Pit	433,2	1938,5	55,9
Kist	449,7	2081,9	85,3
Average	443,9	2087,925	72,725
Standard dev.	26,30551273	146,1649838	12,30782272
Median	441,45	2062,9	74,85
Kies	379,7	2500,3	83,9
Brief	363,5	2289,9	83,3
Dief	308,8	2553,9	118,4
Ziek	291,6	2536,6	81
Average	335,9	2470,175	91,65
Standard dev.	42,3363516	122,2412199	17,8770803
Median	336,15	2518,45	83,6
Put	437,7	1779,1	76,2
Schub	446,2	1636,5	67,8
Kus	379,1	1829,8	124,1
Brug	437,7	1782,9	103,9
Average	425,175	1757,075	93
Standard dev.	30,97691345	83,62473219	25,8411816
Median	437,7	1781	90,05
Muur	321,6	1811,7	118,9
Stuur	343,4	2154,2	155,1
Vuur	300,2	2167,3	156,2
Duur	318,1	1802,9	156,6
Average	320,825	1984,025	146,7
Standard dev.	17,72989472	204,1661313	18,54418148
Median	319,85	1982,95	155,65

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	478,8	2123,2	89,3
Fist	458,9	2024,9	77,6
Six	498,8	2223,8	108,3
Pig	496,7	2262,7	144,2
Average	483,3	2158,65	104,85
Standard dev.	18,57794391	106,796832	29,12438841
Median	487,75	2173,5	98,8
Key	278,1	2465,9	179,1
Sheep	369,2	2502,2	139,5
Peace	359,7	2652,4	176,6
Leaf	358,7	2540,9	183,2
Average	341,425	2540,35	169,6
Standard dev.	42,48100556	80,73357418	20,25026749
Median	359,2	2521,55	177,85
Hug	453,8	1742,5	108,4
Duck	448,7	1700,9	94,3
Cup	470	1710,3	98,7
Bus	646,2	1648,7	103,6
Average	504,675	1700,6	101,25
Standard dev.	94,78598261	38,91614918	6,095079983
Median	461,9	1705,6	101,15
Book	440,8	1158,8	96,8
Foot	394,5	1675	108,5
Hook	388,6	1016,8	80
Cook	376,1	1129,7	91,9
Average	400	1245,075	94,3
Standard dev.	28,26104504	293,0868742	11,80593071
Median	391,55	1144,25	94,35

Ghent_Participant_6

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	489,3	2216,4	91
Sik	522,7	2326,4	79
Pit	521	2090,3	62,4
Kist	505,1	2041,9	92,8
Average	509,525	2168,75	81,3
Standard dev.	15,64062552	128,2827216	14,00999643
Median	513,05	2153,35	85
Kies	400,1	2775,7	74,3
Brief	339,3	2499,7	90,2
Dief	358,1	2526,8	97
Ziek	371,9	2698,5	97,9
Average	367,35	2625,175	89,85
Standard dev.	25,59707014	133,4865131	10,92169096
Median	365	2612,65	93,6
Put	401,9	1814,1	71,2
Schub	509,7	1767,4	81,5
Kus	415,8	2004,1	99,1
Brug	462,2	1822,7	76,2
Average	447,4	1852,075	82
Standard dev.	48,88469426	104,2216348	12,15099447
Median	439	1818,4	78,85
Muur	331,3	2426,1	160
Stuur	359,3	2354,2	125,9
Vuur	329,7	2281,5	180,6
Duur	348,5	2465,3	198,6
Average	342,2	2381,775	166,275
Standard dev.	14,22626679	81,15127335	31,19630053
Median	339,9	2390,15	170,3

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	386,6	2022,7	73,2
Fist	509,9	2050,8	100,2
Six	461,1	2463,7	71,4
Pig	466,6	2052,7	101,7
Average	456,05	2147,475	86,625
Standard dev.	51,18570764	211,262402	16,56872053
Median	463,85	2051,75	86,7
Key	295,1	1424,6	191,9
Sheep	352,8	2793,8	115,5
Peace	283,4	2131,5	151,4
Leaf	327,8	2757,5	170,5
Average	314,775	2276,85	157,325
Standard dev.	31,55443709	644,3907743	32,42143067
Median	311,45	2444,5	160,95
Hug	740,7	1644,8	112,5
Duck	720,9	1693,1	113,1
Cup	679,2	1586	84,6
Bus	735,6	1575,8	97,5
Average	719,1	1624,925	101,925
Standard dev.	27,89300988	54,68475565	13,61919601
Median	728,25	1615,4	105
Book	463,3	1751,1	79,9
Foot	385	1836,9	86,2
Hook	469,6	1427,1	91,8
Cook	320,8	1153,3	109,5
Average	409,675	1542,1	91,85
Standard dev.	70,65	313,5700241	12,73119528
Median	424,15	1589,1	89

Ghent_Participant_7

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	499,6	1974,5	83,8
Sik	492,3	2406,9	85,9
Pit	529,7	2039,1	79,3
Kist	426,1	2073,8	86,1
Average	486,925	2123,575	83,775
Standard dev.	43,66130056	193,3132066	3,159509456
Median	495,95	2056,45	84,85
Kies	295,5	2647,1	78,2
Brief	291,3	2650,5	75,8
Dief	342,9	2606,8	89,1
Ziek	293,1	2637,9	88,7
Average	305,7	2635,575	82,95
Standard dev.	24,85960579	19,90801765	6,941901757
Median	294,3	2642,5	83,45
Put	468,5	1910,7	66,5
Schub	477	1670,9	99
Kus	447,9	1756,9	91,6
Brug	376,3	1734,9	87,1
Average	442,425	1768,35	86,05
Standard dev.	45,74504527	101,669448	13,92611456
Median	458,2	1745,9	89,35
Muur	277	2268,2	180,4
Stuur	316	1922,8	191
Vuur	300	1954,1	184,8
Duur	254,4	2337	207,4
Average	286,85	2120,525	190,9
Standard dev.	26,91139783	212,4945391	11,82821486
Median	288,5	2111,15	187,9

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	428,9	2288	80,7
Fist	506,1	2155,9	106,1
Six	454,5	2528,4	92,6
Pig	470,5	2517,6	144,2
Average	465	2372,475	105,9
Standard dev.	32,31573404	182,039032	27,56120462
Median	462,5	2402,8	99,35
Key	263,6	2885,9	186
Sheep	393,9	2686,9	127,8
Peace	253,8	2724,9	173,3
Leaf	328,9	2786,2	158,7
Average	310,05	2770,975	161,45
Standard dev.	65,08407383	86,85441363	25,05334309
Median	296,25	2755,55	166
Hug	944,3	1398	163,9
Duck	792,3	1537,4	127,4
Cup	738,7	1302,6	107,8
Bus	815,1	1450	113,5
Average	822,6	1422	128,15
Standard dev.	87,22469069	98,20644921	25,21487656
Median	803,7	1424	120,45
Book	428,2	1037	107,9
Foot	443,7	1179,3	123,9
Hook	445,8	1042,8	100,6
Cook	476,4	987,9	88,2
Average	448,525	1061,75	105,15
Standard dev.	20,17281587	82,14517636	14,91229917
Median	444,75	1039,9	104,25

Ghent_Participant_8

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	439,1	1906,1	80,9
Sik	414,9	2057,3	82
Pit	411,6	1769,5	81,9
Kist	410,9	1879,4	93,3
Average	419,125	1903,075	84,525
Standard dev.	13,43040704	118,5991674	5,871044768
Median	413,25	1892,75	81,95
Kies	334,4	2283,2	70,4
Brief	355,5	2071,9	81,6
Dief	333,3	2201,3	87,9
Ziek	299,9	2437,6	75,9
Average	330,775	2248,5	78,95
Standard dev.	22,9790303	153,1647261	7,51731335
Median	333,85	2242,25	78,75
Put	380,9	1534,9	76,4
Schub	428,3	1651,5	80,7
Kus	416,8	1587,4	86,1
Brug	425,4	1825,2	89,9
Average	412,85	1649,75	83,275
Standard dev.	21,85261235	126,3115329	5,937662278
Median	421,1	1619,45	83,4
Muur	348,5	2159,4	137,5
Stuur	332	2055	128,9
Vuur	338,4	2224,1	141,2
Duur	349,1	2156,9	151,4
Average	342	2148,85	139,75
Standard dev.	8,278888814	69,87247908	9,320407716
Median	343,45	2158,15	139,35

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	428	1807,1	71,4
Fist	436,4	1830,1	86,8
Six	434,3	2099,3	76,6
Pig	435,1	2061,1	98,3
Average	433,45	1949,4	83,275
Standard dev.	3,734969879	152,1278848	11,88482927
Median	434,7	1945,6	81,7
Key	321,4	2374,9	193,7
Sheep	259,2	2408,8	125,8
Peace	328,1	2443,4	165,8
Leaf	339,8	2362,3	146,6
Average	312,125	2397,35	157,975
Standard dev.	36,09333964	36,44269474	28,87979397
Median	324,75	2391,85	156,2
Hug	593,1	1316,1	81,2
Duck	656,9	1405,3	97,7
Cup	610	1238,3	89,1
Bus	531,3	1397,8	85,7
Average	597,825	1339,375	88,425
Standard dev.	51,91655966	78,56519055	6,978717647
Median	601,55	1356,95	87,4
Book	442,3	1266,6	63,3
Foot	406,6	1441,9	70,6
Hook	400,9	1317	76,6
Cook	371	1417,1	65,8
Average	405,2	1360,65	69,075
Standard dev.	29,2489316	82,74235111	5,860247435
Median	403,75	1367,05	68,2

Ghent_Participant_9

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	511,8	2234	74,9
Sik	490,3	2326,8	86,6
Pit	491,6	2132,7	81,2
Kist	444,5	2188,4	86
Average	484,55	2220,475	82,175
Standard dev.	28,4565751	82,09991372	5,41871756
Median	490,95	2211,2	83,6
Kies	264,5	2694,1	91,4
Brief	354,7	2673,5	95,4
Dief	350,3	2688,3	103,7
Ziek	276,3	2883,2	64,8
Average	311,45	2734,775	88,825
Standard dev.	47,67847173	99,32939729	16,81593986
Median	313,3	2691,2	93,4
Put	433,6	1887,8	90,4
Schub	417,9	1868,8	92,8
Kus	434,1	2080,2	90,9
Brug	442,8	2017,1	89,8
Average	432,1	1963,475	90,975
Standard dev.	10,36629153	101,9647447	1,297112177
Median	433,85	1952,45	90,65
Muur	305	2357,4	178
Stuur	335	2378,7	185
Vuur	319,2	2534,8	178,2
Duur	343,3	2332,1	196,4
Average	325,625	2400,75	184,4
Standard dev.	16,99968137	91,3740481	8,636357257
Median	327,1	2368,05	181,6

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	538,9	1973,4	83,9
Fist	544,9	2162,3	89,6
Six	441,5	2513,6	89,5
Pig	528,3	2399,2	116,9
Average	513,4	2262,125	94,975
Standard dev.	48,42217123	241,7687514	14,85740556
Median	533,6	2280,75	89,55
Key	404,3	2724,1	167,2
Sheep	385,3	2831,9	132,6
Peace	363,8	2683,1	156,7
Leaf	357	2694,7	188,1
Average	377,6	2733,45	161,15
Standard dev.	21,50178287	67,86385882	23,07820039
Median	374,55	2709,4	161,95
Hug	625,8	1678,4	130,6
Duck	593,3	1717,6	142,5
Cup	664,2	1560,5	125,4
Bus	631,1	1548,4	131,9
Average	628,6	1626,225	132,6
Standard dev.	29,02608023	84,55398177	7,172633175
Median	628,45	1619,45	131,25
Book	416,6	1005,5	129,3
Foot	410,7	1265	148,7
Hook	446,8	1402,8	104,8
Cook	462,1	1130,2	108,6
Average	434,05	1200,875	122,85
Standard dev.	24,48870488	171,3204283	20,31985892
Median	431,7	1197,6	118,95

Ghent_Participant_10

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	486,3	1991,4	69,6
Sik	468,9	2068,2	56,9
Pit	489,9	1978,8	78,3
Kist	477,9	1974,2	83,2
Average	480,75	2003,15	72
Standard dev.	9,364293887	43,97192286	11,53111732
Median	482,1	1985,1	73,95
Kies	297	2475,2	76,8
Brief	285,9	2326,1	90,5
Dief	283	2581,9	80,1
Ziek	363,5	2576,7	71,4
Average	307,35	2489,975	79,7
Standard dev.	37,91644322	119,7841774	8,043631021
Median	291,45	2525,95	78,45
Put	479,1	1590,4	69,7
Schub	457	1584	66,1
Kus	421	1691,6	84,3
Brug	455,1	1790,4	80,6
Average	453,05	1664,1	75,175
Standard dev.	23,98339703	97,56303945	8,660783259
Median	456,05	1641	75,15
Muur	304,7	2197,9	176,7
Stuur	325,9	2082,6	174,7
Vuur	356,5	2119,7	174,1
Duur	293,2	2102,1	184,9
Average	320,075	2125,575	177,6
Standard dev.	27,80508047	50,54155881	4,99199359
Median	315,3	2110,9	175,7

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	451,1	1926,7	69,4
Fist	479,2	1784,3	67,9
Six	456,9	2176,7	70,3
Pig	507,6	2222,5	110
Average	473,7	2027,55	79,4
Standard dev.	25,64150282	207,8400908	20,42400548
Median	468,05	2051,7	69,85
Key	306	2672,6	169,6
Sheep	356,6	2499	135,3
Peace	298,6	2550,2	145,4
Leaf	258,9	2555,5	173,9
Average	305,025	2569,325	156,05
Standard dev.	40,12367339	73,41259542	18,67449241
Median	302,3	2552,85	157,5
Hug	701,1	1499,7	107,2
Duck	736,9	1559,3	110,6
Cup	655,6	1369,8	90,7
Bus	691,5	1453,5	95,7
Average	696,275	1470,575	101,05
Standard dev.	33,42118839	79,93236203	9,394501938
Median	696,3	1476,6	101,45
Book	393,8	1058,6	73,9
Foot	437,8	1212,8	78,9
Hook	438,6	1016,2	81,9
Cook	439,2	886,2	61,7
Average	427,35	1043,45	74,1
Standard dev.	22,37401767	134,6341586	8,90093628
Median	438,2	1037,4	76,4

Ghent_Participant_11

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	566,1	2073,2	63,8
Sik	482,1	2283,4	67,3
Pit	511,6	2022,1	62,4
Kist	544,9	2115,9	76,9
Average	526,175	2123,65	67,6
Standard dev.	36,96695597	113,1926529	6,533503399
Median	528,25	2094,55	65,55
Kies	296,4	2633,7	60,2
Brief	316,2	2506,9	60,9
Dief	299	2467,8	78,8
Ziek	274,2	2711,5	67,5
Average	296,45	2579,975	66,85
Standard dev.	17,23977958	112,7066805	8,618777949
Median	297,7	2570,3	64,2
Put	529,3	1880,4	74,1
Schub	417,4	1687,4	65,5
Kus	572,2	1796,2	86,9
Brug	472,4	1681,6	89,4
Average	497,825	1761,4	78,975
Standard dev.	67,42138014	95,24718719	11,20754954
Median	500,85	1741,8	80,5
Muur	307,5	2082,4	139,6
Stuur	379,4	2073,7	138,4
Vuur	326,2	2065	145,2
Duur	345,8	2150,2	136,5
Average	339,725	2092,825	139,925
Standard dev.	30,72668493	38,9040165	3,741100551
Median	336	2078,05	139

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	478,5	1844,3	65
Fist	529,6	2043,1	91,6
Six	515,5	2446,7	80
Pig	484,5	2388,2	90
Average	502,025	2180,575	81,65
Standard dev.	24,51195287	286,3045625	12,22933631
Median	500	2215,65	85
Key	307,1	2712,7	183,8
Sheep	343,2	2532,1	124,4
Peace	289,6	2694,2	126,5
Leaf	316,3	2664,8	144,5
Average	314,05	2650,95	144,8
Standard dev.	22,36731246	81,65122983	27,52053779
Median	311,7	2679,5	135,5
Hug	812,2	1519,8	113,7
Duck	693,2	1536,4	103,1
Cup	755,4	1551,3	101,9
Bus	791,7	1474	120,4
Average	763,125	1520,375	109,775
Standard dev.	52,19855522	33,48695019	8,84811656
Median	773,55	1528,1	108,4
Book	586,9	1187	97,2
Foot	416,2	1163,8	91,4
Hook	487,7	1220,4	102
Cook	450,3	1018,9	100,4
Average	485,275	1147,525	97,75
Standard dev.	73,77471902	88,84125825	4,680099714
Median	469	1175,4	98,8

Ghent_Participant_12

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	463,8	1973,1	59,8
Sik	468,8	2109,4	61
Pit	422	1956,8	63
Kist	436,8	2097,1	76,9
Average	447,85	2034,1	65,175
Standard dev.	22,23833027	80,28154624	7,927326157
Median	450,3	2035,1	62
Kies	284,8	2583,9	81,2
Brief	374,6	2520,2	101,6
Dief	291,9	2509,4	81,5
Ziek	270,3	2466,7	86,2
Average	305,4	2520,05	87,625
Standard dev.	47,00092198	48,43032796	9,593878256
Median	288,35	2514,8	83,85
Put	398,8	1671,8	87,9
Schub	484,4	1587,2	87,8
Kus	415,8	1811,2	94,8
Brug	424,6	1924,4	110,4
Average	430,9	1748,65	95,225
Standard dev.	37,23958467	149,1895774	10,63402558
Median	420,2	1741,5	91,35
Muur	277,6	2096,4	197,6
Stuur	272,1	1819,5	234,2
Vuur	261,8	2023,7	196
Duur	291,7	2006,5	210,6
Average	275,8	1986,525	209,6
Standard dev.	12,45980203	117,9701198	17,65521642
Median	274,85	2015,1	204,1

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	438,1	1997,3	62,8
Fist	382,7	1917,8	66,8
Six	322,7	2295,2	72,2
Pig	421,5	2438,6	95,5
Average	391,25	2162,225	74,325
Standard dev.	51,25813757	245,6347746	14,6326974
Median	402,1	2146,25	69,5
Key	264,4	2683,6	153
Sheep	286,8	2473	99,2
Peace	302,7	2580,5	129,4
Leaf	309,7	2634	128,9
Average	290,9	2592,775	127,625
Standard dev.	20,09759521	90,26887153	22,0351802
Median	294,75	2607,25	129,15
Hug	655,3	1347,5	84,4
Duck	563,9	1408,7	87,2
Cup	724,2	1358,4	88,1
Bus	648,6	1317,4	76,6
Average	648	1358	84,075
Standard dev.	65,65795204	37,98798056	5,22645514
Median	651,95	1352,95	85,8
Book	395,6	1004,3	63,6
Foot	372,6	1358,8	79,6
Hook	360,1	1167,2	99,8
Cook	375,6	1111,8	79,9
Average	375,975	1160,525	80,725
Standard dev.	14,7047328	148,4788509	14,82191508
Median	374,1	1139,5	79,75

2 Antwerp participants

Antwerp_Participant_1

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	345,1	2690,1	91,4
Sik	298,9	2928,9	81,2
Pit	346,7	2694,5	70,2
Kist	325,8	2817,4	92,3
Average	329,125	2782,725	83,775
Standard dev.	22,2762018	113,918871	10,35579548
Median	335,45	2755,95	86,3
Kies	303,3	2917,4	157
Brief	295	2911,5	192,4
Dief	255,9	2790,1	196,2
Ziek	317,9	2878,3	167
Average	293,025	2874,325	178,15
Standard dev.	26,4984748	58,72843576	19,15298062
Median	299,15	2894,9	179,7
Put	314,8	1915,9	101,5
Schub	330,9	1798,9	95,8
Kus	287,8	2095,2	84,6
Brug	308,1	2061,6	96
Average	310,4	1967,9	94,475
Standard dev.	17,84806245	136,930615	7,093365445
Median	311,45	1988,75	95,9
Muur	327,9	2035,6	201,3
Stuur	292,2	2013,5	195,2
Vuur	334,6	2049,8	195,7
Duur	307,1	1998,1	210,9
Average	315,45	2024,25	200,775
Standard dev.	19,42515551	22,95655317	7,294461369
Median	317,5	2024,55	198,5

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	377,7	2405,8	62
Fist	439,9	2458,7	79,7
Six	384,5	2639,9	81,8
Pig	455,6	2611,2	94,2
Average	414,425	2528,9	79,425
Standard dev.	39,10919542	114,2744358	13,26206997
Median	412,2	2534,95	80,75
Key	401	2948,3	192,4
Sheep	320,6	2807,2	156,4
Peace	348,5	2769,7	155,8
Leaf	374,1	2919,6	176,9
Average	361,05	2861,2	170,375
Standard dev.	34,44807687	86,18820492	17,65793023
Median	361,3	2863,4	166,65
Hug	657,7	1697,2	109,4
Duck	566,2	2049,4	107,5
Cup	527	1553,9	85,2
Bus	513,9	1816,9	111,9
Average	566,2	1779,35	103,5
Standard dev.	64,92092626	209,6928945	12,3323423
Median	546,6	1757,05	108,45
Book	441,3	925,5	117,8
Foot	450,1	1064,6	87,7
Hook	453,7	939,7	93,8
Cook	451,4	971,8	75,3
Average	449,125	975,4	93,65
Standard dev.	5,42486559	62,54091994	17,8453542
Median	450,75	955,75	90,75

Antwerp_Participant_2

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	438,7	2321,5	55,9
Sik	495,2	2546,5	66,2
Pit	492,1	2214,7	60,8
Kist	462,2	2407,9	72,7
Average	472,05	2372,65	63,9
Standard dev.	26,75300606	140,2746235	7,218956527
Median	477,15	2364,7	63,5
Kies	286,1	2802,7	113,3
Brief	280,9	2783,9	75;5
Dief	297,7	2718,6	93,3
Ziek	290,8	2723,4	91,2
Average	288,875	2757,15	99,26666667
Standard dev.	7,138802421	42,48737067	12,19849718
Median	288,45	2753,65	93,3
Put	512,4	1797,5	57
Schub	504,4	1735,1	79,2
Kus	473,8	1898,3	87,3
Brug	478,3	1855,1	73,1
Average	492,225	1821,5	74,15
Standard dev.	19,04947506	70,87115069	12,82770439
Median	491,35	1826,3	76,15
Muur	290,7	1989,8	168,2
Stuur	274,2	2321,2	169,9
Vuur	261,4	2272,6	183,9
Duur	265,2	2342,8	191,6
Average	272,875	2231,6	178,4
Standard dev.	13,03926762	163,8513961	11,26617356
Median	269,7	2296,9	176,9

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	443,8	2256,2	71,3
Fist	503,6	2305,5	75,6
Six	516,6	2381,5	82,9
Pig	483,5	2506,2	110,4
Average	486,875	2362,35	85,05
Standard dev.	31,78127069	108,8719278	17,56521183
Median	493,55	2343,5	79,25
Key	358,5	2720,8	206,3
Sheep	319,8	2725,8	108,4
Peace	321,1	2731,5	150,9
Leaf	289,9	2776,5	153,3
Average	322,325	2738,65	154,725
Standard dev.	28,09440929	25,60917804	40,09425354
Median	320,45	2728,65	152,1
Hug	644,7	1607,6	107,9
Duck	795,3	1221,4	128
Cup	811,4	1493,7	96,6
Bus	766	1419,4	139,2
Average	754,35	1435,525	117,925
Standard dev.	75,47725927	162,3819238	19,23007627
Median	780,65	1456,55	117,95
Book	505,7	1173,8	85,3
Foot	543,8	1695,8	86,8
Hook	486,1	1729,9	97,8
Cook	468,3	1341,8	95,4
Average	500,975	1485,325	91,325
Standard dev.	32,3791471	271,8847228	6,199663969
Median	495,9	1518,8	91,1

Antwerp_Participant_3

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	352,6	1989,5	60,6
Sik	347,9	2204,5	81,5
Pit	325,5	2022,5	57,8
Kist	299,7	2109,7	66,2
Average	331,425	2081,55	66,525
Standard dev.	24,23074837	96,3833146	10,57650698
Median	336,7	2066,1	63,4
Kies	274,7	2369,7	120,2
Brief	328,6	2236,9	120,1
Dief	281,1	2292,7	131,3
Ziek	270,8	2251,4	128,8
Average	288,8	2287,675	125,1
Standard dev.	26,87092605	59,57440026	5,806318398
Median	277,9	2272,05	124,5
Put	299,4	1788,7	72,1
Schub	377	1507,4	78,5
Kus	359,6	1732,8	69,3
Brug	397,7	1641,9	78,9
Average	358,425	1667,7	74,7
Standard dev.	42,31976489	122,8016558	4,760952286
Median	368,3	1687,35	75,3
Muur	307,6	1818,7	168,8
Stuur	327,4	1900,2	155,2
Vuur	272,6	1875,9	145,6
Duur	289,3	1984,9	127
Average	299,225	1894,925	149,15
Standard dev.	23,60344255	69,03025303	17,56843761
Median	298,45	1888,05	150,4

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	431,7	1753,3	79,6
Fist	377,9	1726,9	63,9
Six	390,4	1879,8	86,4
Pig	369;6	2106,4	90,1
Average	400	1866,6	80
Standard dev.	28,15546128	173,2351581	11,58073112
Median	390,4	1816,55	83
Key	281,3	2248,2	178,7
Sheep	316,9	2297,7	87,5
Peace	292,3	2287,1	165,6
Leaf	281,9	2189,5	127,2
Average	293,1	2255,625	139,75
Standard dev.	16,6509259	48,95108954	41,12132456
Median	287,1	2267,65	146,4
Hug	514,7	1269,5	108,1
Duck	531,1	1353,3	95,8
Cup	550,4	1180,6	79,9
Bus	512,3	1224,1	96,3
Average	527,125	1256,875	95,025
Standard dev.	17,62278355	73,82241642	11,57508099
Median	522,9	1246,8	96,05
Book	413,7	1077,3	62,7
Foot	423,1	1136,1	70,4
Hook	385,1	1020,7	72,4
Cook	385,8	1042	49
Average	401,925	1069,025	63,625
Standard dev.	19,40899705	50,44200465	10,6089192
Median	399,75	1059,65	66,55

Antwerp_Participant_4

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	418,4	2437,8	73,5
Sik	436,8	2466,6	63,8
Pit	386,6	2355,3	62,7
Kist	409,3	2501,3	68,5
Average	412,775	2440,25	67,125
Standard dev.	20,86502257	62,30016051	4,93853892
Median	413,85	2452,2	66,15
Kies	314,3	2645,1	112,4
Brief	370,6	2690,2	142,8
Dief	270,1	2702,6	132,9
Ziek	340,8	2720,2	102,5
Average	323,95	2689,525	122,65
Standard dev.	42,63445399	32,07256096	18,45869262
Median	327,55	2696,4	122,65
Put	427,3	2334,9	74,3
Schub	370,9	2178,1	72,7
Kus	373,6	2307,3	80,2
Brug	396,2	2342,6	86,4
Average	392	2290,725	78,4
Standard dev.	26,12470096	76,59757938	6,232709416
Median	384,9	2321,1	77,25
Muur	331,4	2354,6	162,4
Stuur	362,9	2526,4	132,3
Vuur	299,4	2365,5	171,5
Duur	349,2	2428,5	176,7
Average	335,725	2418,75	160,725
Standard dev.	27,43639614	78,81279507	19,85016793
Median	340,3	2397	166,95

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	450,7	2181,2	69,2
Fist	396,9	2215,4	85,1
Six	454,1	2359,8	77,8
Pig	483,5	2396,9	77,9
Average	446,3	2288,325	77,5
Standard dev.	36,07584603	105,9733103	6,503845017
Median	452,4	2287,6	77,85
Key	362,3	2866,5	156,3
Sheep	273,1	2715,8	107,1
Peace	285,7	2568,1	136,8
Leaf	255,5	2833,1	148,5
Average	294,15	2745,875	137,175
Standard dev.	47,09122353	134,9901817	21,59218609
Median	279,4	2774,45	142,65
Hug	719,4	1685,7	100,2
Duck	618,4	1656,9	95,1
Cup	704	1569,2	86,2
Bus	643,3	1724,4	96,7
Average	671,275	1659,05	94,55
Standard dev.	48,18473998	65,97598048	5,960145412
Median	673,65	1671,3	95,9
Book	487,5	1032,7	92,2
Foot	476,1	1387,6	86,4
Hook	470,4	1067,2	66,4
Cook	518	1095,7	90,5
Average	488	1145,8	83,875
Standard dev.	21,22592754	163,2450306	11,90164554
Median	481,8	1081,45	88,45

Antwerp_Participant_5

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	358	2313,6	70,5
Sik	370,7	2395,8	83,1
Pit	369,1	2292,7	57,2
Kist	346,1	2363,8	61,2
Average	360,975	2341,475	68
Standard dev.	11,41209154	46,92414979	11,50565079
Median	363,55	2338,7	65,85
Kies	354,9	2567,8	139,4
Brief	305,7	2641,6	123
Dief	287,3	2497,9	121,1
Ziek	370,9	2655,8	108,3
Average	329,7	2590,775	122,95
Standard dev.	39,60740671	72,94986292	12,7625755
Median	330,3	2604,7	122,05
Put	366,2	1808,1	62,1
Schub	355,7	1819	68,8
Kus	301,7	2058,2	74
Brug	327,3	1821,4	68,4
Average	337,725	1876,675	68,325
Standard dev.	29,09987113	121,1549799	4,871259249
Median	341,5	1820,2	68,6
Muur	292,9	1865,6	165,3
Stuur	378,2	2093,9	163,5
Vuur	323,6	2036,1	182,1
Duur	315,8	2106,8	153,5
Average	327,625	2025,6	166,1
Standard dev.	36,14640369	111,0083781	11,8625461
Median	319,7	2065	164,4

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	349,7	2112,7	55,1
Fist	367,6	2310,1	69,1
Six	399,6	2400,5	62,8
Pig	341,3	2482,9	79,6
Average	364,55	2326,55	66,65
Standard dev.	25,81272296	159,0768682	10,35905401
Median	358,65	2355,3	65,95
Key	366,1	2594,5	146,5
Sheep	334,5	2618,7	98,3
Peace	327,5	2574,7	124,7
Leaf	336,8	2528,3	127,9
Average	341,225	2579,05	124,35
Standard dev.	17,0484359	38,32018615	19,84901341
Median	335,65	2584,6	126,3
Hug	486,9	1387,3	98,8
Duck	577,1	1495,7	96,7
Cup	598,3	1458,6	58,7
Bus	599,8	1331,9	99,9
Average	565,525	1418,375	88,525
Standard dev.	53,43172435	73,12269939	19,92759812
Median	587,7	1422,95	97,75
Book	367,6	999,6	73,7
Foot	393,2	1254,8	85,5
Hook	402,8	1156,9	74,7
Cook	359,6	994,9	68,9
Average	380,8	1101,55	75,7
Standard dev.	20,50560899	126,908379	7,006663495
Median	380,4	1078,25	74,2

Antwerp_Participant_6

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	348,9	2305,5	57,3
Sik	475,5	2434,1	64,9
Pit	454,6	2435,6	71
Kist	401,5	2315,5	57,8
Average	420,125	2372,675	62,75
Standard dev.	56,7885185	71,91209333	6,503588753
Median	428,05	2374,8	61,35
Kies	282,1	2334,9	123,7
Brief	309,8	2620,1	122,2
Dief	325,3	2415,7	145,1
Ziek	313,5	2439,4	100,1
Average	307,675	2452,525	122,775
Standard dev.	18,28631091	120,3403057	18,38248714
Median	311,65	2427,55	122,95
Put	363,4	2174,1	76,6
Schub	390,6	1994,9	81,6
Kus	378,7	1844,5	72,2
Brug	416,3	1980,3	77,8
Average	387,25	1998,45	77,05
Standard dev.	22,33868095	135,2720099	3,872552991
Median	384,65	1987,6	77,2
Muur	314,8	2227,1	130,8
Stuur	315,1	2128,8	126,1
Vuur	336,2	2152	144,4
Duur	301,7	2252,2	165,9
Average	316,95	2190,025	141,8
Standard dev.	14,2731683	58,97651369	17,84245872
Median	314,95	2189,55	137,6

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	465,2	2103,7	77,6
Fist	432,6	2225,3	69,8
Six	474,8	2348,3	87,6
Pig	453	2415,7	109,1
Average	456,4	2273,25	86,025
Standard dev.	18,20256392	137,8053095	17,02123674
Median	459,1	2286,8	82,6
Key	348,2	2607,9	175,3
Sheep	334,7	2483,9	116
Peace	316,2	2651,2	143,4
Leaf	303,8	2605,9	131,6
Average	325,725	2587,225	141,575
Standard dev.	19,63931007	71,98395076	25,12825435
Median	325,45	2606,9	137,5
Hug	504,2	1816,1	115,5
Duck	542,8	1748,9	98
Cup	503,1	1525,1	87,1
Bus	595,2	1619,5	99,8
Average	536,325	1677,4	100,1
Standard dev.	43,37475264	130,2531382	11,69985755
Median	523,5	1684,2	98,9
Book	451,5	1047,1	89,9
Foot	450,6	1208,6	77,3
Hook	443,2	1090,1	71,7
Cook	485	1150,7	84,8
Average	457,575	1124,125	80,925
Standard dev.	18,65768385	70,55212612	8,03756804
Median	451,05	1120,4	81,05

Antwerp_Participant_7

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	437,3	2543,7	50,2
Sik	388,4	2556,6	78
Pit	412,2	2592,1	71,4
Kist	321,9	2593,5	71,7
Average	389,95	2571,475	67,825
Standard dev.	49,56574758	25,18734801	12,13764804
Median	400,3	2574,35	71,55
Kies	305,9	2661,7	153,4
Brief	322,8	2647,9	179,6
Dief	336,5	2665	175,1
Ziek	414,9	2653,6	177,4
Average	345,025	2657,05	171,375
Standard dev.	48,23521362	7,755643107	12,123359
Median	329,65	2657,65	176,25
Put	413,1	1985,9	80,2
Schub	407,6	2081,8	74,8
Kus	427,7	2158,4	79,9
Brug	414,6	2054,8	97,1
Average	415,75	2070,225	83
Standard dev.	8,516063253	71,31490611	9,721111048
Median	413,85	2068,3	80,05
Muur	450,8	1982,5	210,9
Stuur	378,9	2047,4	204,5
Vuur	362,6	2147,3	213,1
Duur	377,9	2097,1	203,5
Average	392,55	2068,575	208
Standard dev.	39,54326744	70,4002545	4,722993401
Median	378,4	2072,25	207,7

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	460,5	2505,9	79,1
Fist	482,4	2455,2	83,2
Six	358,3	2575,9	76,4
Pig	377,2	2546,4	102,1
Average	419,6	2520,85	85,2
Standard dev.	61,02485832	52,33459659	11,60833034
Median	418,85	2526,15	81,15
Key	333,9	2656,8	197,9
Sheep	432,4	2648,2	195,9
Peace	367,4	2584,2	156,2
Leaf	383,7	2702,5	204
Average	379,35	2647,925	188,5
Standard dev.	40,99483707	48,71053103	21,80718536
Median	375,55	2652,5	196,9
Hug	690,1	1689,4	127,5
Duck	735,4	1366,6	182,8
Cup	420,2	1369,3	109,9
Bus	645,7	1680,7	137,4
Average	622,85	1526,5	139,4
Standard dev.	139,9752478	183,1155373	31,08815423
Median	667,9	1525	132,45
Book	442,4	917,8	123,2
Foot	387,9	917,2	108,1
Hook	480,9	1103,8	91,8
Cook	493,6	1082	80,9
Average	451,2	1005,2	101
Standard dev.	47,4839622	101,6578575	18,54633836
Median	461,65	999,9	99,95

Antwerp_Participant_8

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	471,6	2121,6	62,7
Sik	450,4	2210,3	69,4
Pit	466	1940,8	58,8
Kist	435,7	2116,4	81,7
Average	455,925	2097,275	68,15
Standard dev.	16,19452068	112,8665104	10,03809411
Median	458,2	2119	66,05
Kies	281,8	2460,6	129,8
Brief	311,9	2522,6	138,6
Dief	273,6	2515,8	132,7
Ziek	289,2	2527,2	120,9
Average	289,125	2506,55	130,5
Standard dev.	16,46600032	30,98919166	7,373375527
Median	285,5	2519,2	131,25
Put	352,6	1681,8	63
Schub	456,1	1660,3	79,7
Kus	444,9	1784	68,3
Brug	463,7	1701,5	69,3
Average	429,325	1706,9	70,075
Standard dev.	51,72957729	54,08370056	6,986832854
Median	450,5	1691,65	68,8
Muur	326,3	1900,1	136,4
Stuur	289,9	2048,8	155,1
Vuur	275,6	1889,4	136,4
Duur	246,4	2183,7	175,1
Average	284,55	2005,5	150,75
Standard dev.	33,20245975	139,3060181	18,47241186
Median	282,75	1974,45	145,75

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	456	2136,7	63,4
Fist	390,8	2149	84
Six	491,6	2284,7	69,7
Pig	470,8	2224,4	89,2
Average	452,3	2198,7	76,575
Standard dev.	43,52271438	69,2112226	12,04668004
Median	463,4	2186,7	76,85
Key	305,9	2640,3	158,8
Sheep	296,3	2512,9	115,9
Peace	381,6	2543,4	157,9
Leaf	302,6	2481,7	161,7
Average	321,6	2544,575	148,575
Standard dev.	40,19776113	68,60813241	21,84359174
Median	304,25	2528,15	158,35
Hug	607,3	1476,6	92,2
Duck	636,4	1505,4	126
Cup	536	1323,2	91,8
Bus	574,5	1546,6	124,4
Average	588,55	1462,95	108,6
Standard dev.	43,20528517	97,49480328	19,17985054
Median	590,9	1491	108,3
Book	447,7	1024,4	82,7
Foot	434	1219,8	89,9
Hook	423,9	1000,3	93
Cook	484,1	1080,3	87,1
Average	447,425	1081,2	88,175
Standard dev.	26,32354776	98,28838521	4,37369028
Median	440,85	1052,35	88,5

Antwerp_Participant_9

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	410,7	2464,7	83,4
Sik	404,5	2383,3	55,6
Pit	415,5	2466,3	71,9
Kist	364,5	2447,7	81,9
Average	398,8	2440,5	73,2
Standard dev.	23,30579327	39,05107766	12,79557215
Median	407,6	2456,2	76,9
Kies	364	2739,2	145,5
Brief	363,8	2702,5	132,9
Dief	330	2672,8	141
Ziek	311,9	2765,4	117,1
Average	342,425	2719,975	134,125
Standard dev.	25,8748752	40,67713322	12,49009608
Median	346,9	2720,85	136,95
Put	408,5	1992,3	71,4
Schub	414,9	2078,8	73,6
Kus	380,2	1982,9	94,3
Brug	372,6	2362,9	83,9
Average	394,05	2104,225	80,8
Standard dev.	20,78019891	177,769652	10,5207731
Median	394,35	2035,55	78,75
Muur	296,4	2362,6	168,6
Stuur	371,9	2541,6	162,6
Vuur	286,5	2369	178,8
Duur	288,5	2391,7	184,9
Average	310,825	2416,225	173,725
Standard dev.	40,94039366	84,51076361	10,01111882
Median	292,45	2380,35	173,7

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	446,3	2166,5	73
Fist	536,8	2141,9	82,4
Six	455,3	2483,3	78
Pig	472,2	2586,2	95,8
Average	477,65	2344,475	82,3
Standard dev.	40,86877374	223,915972	9,785022569
Median	463,75	2324,9	80,2
Key	355,4	2837,4	180,8
Sheep	333,6	2725,7	130,8
Peace	293,7	2813,7	155,1
Leaf	353,6	2733,5	170,1
Average	334,075	2777,575	159,2
Standard dev.	28,67256238	56,3254457	21,66979465
Median	343,6	2773,6	162,6
Hug	681,4	1579,9	133,9
Duck	727,6	1776,2	148,6
Cup	681,7	1731,6	104,4
Bus	677,8	1653,2	108,2
Average	692,125	1685,225	123,775
Standard dev.	23,71629187	86,69111354	21,10898308
Median	681,55	1692,4	121,05
Book	455,9	988,3	99,5
Foot	486,2	1203,5	95,4
Hook	481,4	1232,2	127,1
Cook	538,1	1377,1	92,6
Average	490,4	1200,275	103,65
Standard dev.	34,46824626	160,4473823	15,88804582
Median	483,8	1217,85	97,45

Antwerp_Participant_10

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	475,8	2389,4	93,1
Sik	406,7	2599,7	60,6
Pit	349,7	2482,3	60,5
Kist	503,6	2362,8	80,9
Average	433,95	2458,55	73,775
Standard dev.	69,38580546	107,1403597	16,06266375
Median	441,25	2435,85	70,75
Kies	361,3	2374,1	121,6
Brief	419,7	2780,3	135,9
Dief	371,9	2266,2	153,4
Ziek	357,5	2502,5	150,5
Average	377,6	2480,775	140,35
Standard dev.	28,7204921	221,8176334	14,65935424
Median	366,6	2438,3	143,2
Put	373,9	2064,2	56
Schub	427,4	2137,2	65,3
Kus	388,5	1956,5	132,7
Brug	419,3	1952,4	109,9
Average	402,275	2027,575	90,975
Standard dev.	25,27190469	89,55811428	36,43069905
Median	403,9	2010,35	87,6
Muur	520,4	2151,5	147,9
Stuur	470,4	1902,4	204,8
Vuur	439,3	1956,1	228,1
Duur	419,1	2064,2	173,9
Average	462,3	2018,55	188,675
Standard dev.	44,10767129	111,2811005	35,09628423
Median	454,85	2010,15	189,35

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	450,7	2240,7	51,4
Fist	396,6	2429,4	77,2
Six	432,9	2374,6	91,3
Pig	399,4	2521,6	98,3
Average	419,9	2391,575	79,55
Standard dev.	26,33616019	117,4572937	20,71690131
Median	416,15	2402	84,25
Key	427,9	2692,3	180,7
Sheep	367	2044,9	134
Peace	397	2214,8	162,2
Leaf	458,5	2663,1	146,3
Average	412,6	2403,775	155,8
Standard dev.	39,42765527	324,0364936	20,21929771
Median	412,45	2438,95	154,25
Hug	640,2	1572,1	115,3
Duck	575,4	1626,6	118,9
Cup	633,4	1512,1	90,8
Bus	623,1	1483,9	111
Average	618,025	1548,675	109
Standard dev.	29,27323863	63,65157107	12,55574238
Median	628,25	1542,1	113,15
Book	444,8	1084,6	98,3
Foot	420,9	1008,8	113,5
Hook	445,6	1132,9	93,2
Cook	466,2	1101,6	95,3
Average	444,375	1081,975	100,075
Standard dev.	18,52104659	52,72560257	9,191436232
Median	445,2	1093,1	96,8

Antwerp_Participant_11

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	413,3	2462,4	56,1
Sik	315,5	2560,5	72,9
Pit	291,4	2589,6	55
Kist	354,1	2545	77,1
Average	343,575	2539,375	65,275
Standard dev.	53,175143	54,54541686	11,36848715
Median	334,8	2552,75	64,5
Kies	454,9	2530,2	160,3
Brief	323,1	2626,8	169,8
Dief	339,7	2608,2	181,6
Ziek	400,2	2588,8	168,4
Average	379,475	2588,5	170,025
Standard dev.	60,21806899	41,84877537	8,779664003
Median	369,95	2598,5	169,1
Put	322,4	2010,8	69,7
Schub	393,7	2039,3	72,7
Kus	326,4	2199,2	103,1
Brug	397,5	2147,1	87,9
Average	360	2099,1	83,35
Standard dev.	41,16900128	88,87620604	15,38949858
Median	360,05	2093,2	80,3
Muur	379,1	2067,6	213,9
Stuur	379,8	1983,9	221,7
Vuur	356,1	2136,5	219,2
Duur	338,5	2063,3	198
Average	363,375	2062,825	213,2
Standard dev.	19,90600831	62,39708193	10,64236816
Median	367,6	2065,45	216,55

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	445,1	2315,6	57,3
Fist	367,4	2385,1	75,4
Six	461,3	2389,2	79
Pig	436,7	2559,7	95,2
Average	427,625	2412,4	76,725
Standard dev.	41,42779864	103,8445312	15,55192914
Median	440,9	2387,15	77,2
Key	409,9	2637,4	232,4
Sheep	319,3	2566,3	145,9
Peace	373,1	2564,4	189,4
Leaf	363,6	2649	199,1
Average	366,475	2604,275	191,7
Standard dev.	37,25304238	45,20216625	35,65660294
Median	368,35	2601,85	194,25
Hug	728,6	1603,9	112,8
Duck	687,9	1653,7	95,3
Cup	738,7	1690,8	102,5
Bus	645,1	1467,1	148,9
Average	700,075	1603,875	114,875
Standard dev.	42,72410522	97,8874992	23,7930487
Median	708,25	1628,8	107,65
Book	189,5	1066,3	115,7
Foot	413,6	1108,2	86,1
Hook	443,4	966,4	108,4
Cook	452,7	1115,9	97,1
Average	374,8	1064,2	101,825
Standard dev.	124,6540546	68,74624838	12,97879681
Median	428,5	1087,25	102,75

Antwerp_Participant_12

DUTCH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Schip	330,6	2403,5	58,7
Sik	370,1	2428,5	64,6
Pit	305,9	2421,4	66,6
Kist	344,1	2428,4	73
Average	337,675	2420,45	65,725
Standard dev.	26,78536603	11,77865301	5,896538533
Median	337,35	2424,9	65,6
Kies	334,8	2408,4	115,4
Brief	308	2616,6	120,9
Dief	297,5	2371,2	129,2
Ziek	356,9	2424,8	121,8
Average	324,3	2455,25	121,825
Standard dev.	26,81380242	109,8792519	5,672374576
Median	321,4	2416,6	121,35
Put	302,7	1776,1	52,5
Schub	368,4	1852,4	65,3
Kus	389,2	1831,3	80,5
Brug	396,2	1562,3	63,6
Average	364,125	1755,525	65,475
Standard dev.	42,61794419	132,7727978	11,5129999
Median	378,8	1803,7	64,45
Muur	340,7	1753,9	174,3
Stuur	313,1	1855,9	176,1
Vuur	369,8	1994,2	194,1
Duur	318,9	1998,6	192
Average	335,625	1900,65	184,125
Standard dev.	25,69557355	118,1579875	10,36737672
Median	329,8	1925,05	184,05

ENGLISH			
Stimulus	F1 (Hz)	F2 (Hz)	Duration (ms)
Spit	364,8	2185,9	78,3
Fist	371,3	2367,3	55,9
Six	400,2	2401,9	62,3
Pig	361,6	2391,2	86,8
Average	374,475	2336,575	70,825
Standard dev.	17,61843258	101,4860048	14,21862042
Median	368,05	2379,25	70,3
Key	349	2453,9	214,5
Sheep	297,1	2393,3	108,7
Peace	348,1	2464,4	145,6
Leaf	294,2	2539,7	145,1
Average	322,1	2462,825	153,475
Standard dev.	30,56697564	60,07114532	44,20032994
Median	322,6	2459,15	145,35
Hug	597,6	1541,9	106,7
Duck	427	1558,6	87,3
Cup	507,8	1522,7	83,3
Bus	518,5	1433,8	93,8
Average	512,725	1514,25	92,775
Standard dev.	69,78573756	55,6029076	10,24219215
Median	513,15	1532,3	90,55
Book	401,9	876,6	88,8
Foot	376,1	1403,8	82,8
Hook	384,4	865,8	70,8
Cook	415,5	1136,2	72,8
Average	394,475	1070,6	78,8
Standard dev.	17,66661164	254,8886293	8,485281374
Median	393,15	1006,4	77,8

Appendix G. Production experiment: average values (stimuli)

1 Dutch production task

1.1 Ghent participants

kies/k i s/ 'molar'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	265,575	2192,525	74,725
	Standard deviation	48,70074435	142,6596737	11,59062121
	Median	254,15	2194,4	74,45
Female (n=8)	Average	314,8375	2616,35	76,3625
	Standard deviation	48,02511062	97,94152775	10,04830298
	Median	296,7	2627,25	77,5

brief /b r i f/'letter'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	318,65	2165,55	79,65
	Standard deviation	29,7124553	125,7934418	7,622991539
	Median	318	2125,25	79,75
Female (n=8)	Average	327,975	2518,8125	86,3875
	Standard deviation	34,66784224	150,4782228	12,89799957
	Median	327,75	2513,55	90,35

dief /d i f/ 'thief'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	274,75	2076,1	75,775
	Standard deviation	41,78432721	112,2381694	10,7434864
	Median	265,7	2084,95	76,55
Female (n=8)	Average	312,7125	2590,075	90,7875
	Standard deviation	33,65707047	103,5345594	14,57046303
	Median	303,9	2567,9	85,3

ziek /z i k/ 'ill'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	298,875	2281,675	70,525
	Standard deviation	10,84077949	157,0658752	7,491495178
	Median	301,3	2281,5	71,15
Female (n=8)	Average	306,6	2670,175	80,1625
	Standard deviation	40,04030113	146,0715749	11,40813463
	Median	292,35	2668,2	82,4

schip /s x ɪ p/ 'ship'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	431,45	1779,075	70,325
	Standard deviation	34,99738085	98,33166242	13,35524242
	Median	432,75	1767	71,1
Female (n=8)	Average	496,8	2083,8375	70,2375
	Standard deviation	31,50201808	107,6038095	14,16514207
	Median	487,8	2058,55	71,5

sik /s ɪ k/ ‘goatee’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	414,35	2008,25	69
	Standard deviation	33,84321695	35,36971397	9,826155572
	Median	419,75	2000,65	67,35
Female (n=8)	Average	470,125	2267,275	70,8875
	Standard deviation	36,38487716	116,8952003	12,8478389
	Median	475,5	2306,9	71,8

pit /p ɪ t/‘kernel’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	417,2	1726,475	66,5
	Standard deviation	17,2333011	55,79022465	10,57827963
	Median	415,6	1724,3	63,15
Female (n=8)	Average	484,6	2018,55	67,225
	Standard deviation	39,22830611	66,78721007	10,68132549
	Median	490,75	2006,1	62,7

kist /k ɪ s t/ ‘chest’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	433,55	1792,5	80,175
	Standard deviation	22,94420769	69,05616072	15,83190344
	Median	429,7	1781,85	84,25
Female (n=8)	Average	470,375	2091,05	81,5
	Standard deviation	39,74758753	66,22181341	8,515867542
	Median	463,8	2089,5	84,25

muur /m y r/ ‘wall’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	296,05	1919,25	162,625
	Standard deviation	37,50559958	169,565769	24,50732884
	Median	285,2	1860,65	163,35
Female (n=8)	Average	305	2118,6375	164,225
	Standard deviation	19,32592633	251,8540901	25,07740873
	Median	306,25	2147,15	169,65

stuur /s t y r/ ‘steering wheel’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	289,65	2022,5	159,975
	Standard deviation	29,37737224	105,980407	22,69836631
	Median	280,75	2047,1	163,8
Female (n=8)	Average	331,6875	2107,9	170,8875
	Standard deviation	31,85594528	191,3525094	33,84977052
	Median	330,45	2080,05	168,75

vuur /v y r/‘fire’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	273,125	2012,45	167,275
	Standard deviation	49,19616347	175,6139801	30,15353766
	Median	260,65	1980,65	158,55
Female (n=8)	Average	317,0875	2173,675	171,175
	Standard deviation	29,54524073	182,1910359	17,45350968
	Median	322,7	2143,5	176,15

duur /d y r/‘expensive’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	270,225	1980,775	173,1
	Standard deviation	57,93164794	126,6092776	21,00206339
	Median	258,45	1943,2	170,15
Female (n=8)	Average	313,05	2153,9375	179,5875
	Standard deviation	32,89120595	215,1138366	29,13267373
	Median	313,75	2126,15	190,65

put /p ʏ t/ ‘pit’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	420,05	1452,05	71,125
	Standard deviation	51,14642379	60,06094127	5,504770658
	Median	402	1436,9	71,55
Female (n=8)	Average	446,575	1776,0875	76,3875
	Standard deviation	43,80263365	118,7021535	8,495450463
	Median	435,65	1796,6	74,6

schub /s x ʏ p/ ‘scale’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	454,275	1586,475	75,525
	Standard deviation	60,30203838	79,94566384	8,149181963
	Median	435,3	1604,6	78,45
Female (n=8)	Average	455,5125	1683,05	77,1
	Standard deviation	33,01559967	95,15722328	15,2813612
	Median	451,6	1666,55	74,65

kus /k ʏ s/ ‘kiss’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	442,475	1565,45	83,175
	Standard deviation	67,38997824	66,83624266	10,92745014
	Median	415,2	1578,95	81,15
Female (n=8)	Average	440,7375	1846,2375	93,3375
	Standard deviation	57,08489761	129,6708463	14,3925514
	Median	427,55	1805,55	91,25

brug /b r ʏ x/‘bridge’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	414,45	1610,875	88,1
	Standard deviation	16,48726781	149,6153373	11,85017581
	Median	421,25	1569,9	85,9
Female (n=8)	Average	439,2	1806,4125	87,5
	Standard deviation	29,47085728	114,6314148	15,10789765
	Median	442,65	1786,65	88,25

1.2 Antwerp participants

kies/k i s/ 'molar'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	314,8	2468,75	129,8
	Standard deviation	56,70996385	140,0778534	13,5764502
	Median	314,8	2468,75	129,8
Female (n=10)	Average	328,85	2587,43	133,24
	Standard deviation	53,57296686	196,5794499	18,97959138
	Median	310,1	2587,65	126,75

brief /b r i f/'letter'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	317,15	2439,25	121,55
	Standard deviation	16,19274529	286,1661143	2,050609665
	Median	317,15	2439,25	121,55
Female (n=10)	Average	330,56	2690,24	148,3444444
	Standard deviation	41,88431687	110,6808445	25,82141123
	Median	317,35	2669,05	138,6

dief /d i f/ 'thief'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	284,2	2395,3	126,2
	Standard deviation	4,384062043	145,0983115	7,212489168
	Median	284,2	2395,3	126,2
Female (n=10)	Average	309,82	2572,62	148,05
	Standard deviation	36,81943327	172,4510159	29,96439554
	Median	311,5	2636,6	143,05

ziek /z i k/ 'ill'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	320,85	2453,6	118,55
	Standard deviation	70,7813888	285,9539823	14,49568901
	Median	320,85	2453,6	118,55
Female (n=10)	Average	339,36	2622,36	131,69
	Standard deviation	43,21008627	150,4717929	31,54215698
	Median	329,35	2621,2	121,35

schip /s x ɪ p/ 'ship'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	355,3	2151,55	65,55
	Standard deviation	3,818376618	229,1733078	7,000357134
	Median	355,3	2151,55	65,55
Female (n=10)	Average	409,04	2414,02	68,23
	Standard deviation	51,6564334	150,9504908	15,91079647
	Median	415,85	2420,65	60,7

sik /s ɪ k/ 'goatee'		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	359,3	2300,15	82,3
	Standard deviation	16,12203461	135,2695272	1,13137085
	Median	359,3	2300,15	82,3
Female (n=10)	Average	404,2	2511,5	67,72
	Standard deviation	64,06286843	185,5066516	7,817331173
	Median	405,6	2506,55	65,55

pit /p ɪ t/‘kernel’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	347,3	2157,6	57,5
	Standard deviation	30,82985566	191,0602523	0,424264069
	Median	347,3	2157,6	57,5
Female (n=10)	Average	392,07	2419,26	64,89
	Standard deviation	68,13313845	214,9843571	6,107090597
	Median	399,4	2450,95	64,65

kist /k ɪ s t/ ‘chest’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	322,9	2236,75	63,7
	Standard deviation	32,80975465	179,6758331	3,535533906
	Median	322,9	2236,75	63,7
Female (n=10)	Average	392,27	2453,59	75,76
	Standard deviation	60,97779286	184,4903277	9,330380485
	Median	383	2438,05	75,05

muur /m y r/ ‘wall’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	300,25	1842,15	167,05
	Standard deviation	10,39446968	33,16330804	2,474873734
	Median	300,25	1842,15	167,05
Female (n=10)	Average	357,85	2082,53	171,47
	Standard deviation	73,43812588	194,4871321	29,42704916
	Median	329,65	2051,6	168,4

stuur /s t y r/ ‘steering wheel’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	352,8	1997,05	159,35
	Standard deviation	35,92102448	136,9665835	5,868986284
	Median	352,8	1997,05	159,35
Female (n=10)	Average	344,84	2136,99	174,83
	Standard deviation	59,51403569	244,4685913	31,8650296
	Median	339	2048,1	173

vuur /v y r/‘fire’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	298,1	1956	163,85
	Standard deviation	36,06244584	113,2785063	25,80939751
	Median	298,1	1956	163,85
Female (n=10)	Average	332,15	2133,24	186,52
	Standard deviation	53,5223785	165,5428189	30,22742684
	Median	335,4	2141,9	189

duur /d y r/‘expensive’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	302,55	2045,85	140,25
	Standard deviation	18,7383297	86,19631663	18,7383297
	Median	302,55	2045,85	140,25
Female (n=10)	Average	321,25	2182,02	187,25
	Standard deviation	51,92649185	162,6698142	14,44885616
	Median	313	2140,4	188,25

put /p ʏ t/ ‘pit’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	332,8	1798,4	67,1
	Standard deviation	47,23473298	13,71787156	7,071067812
	Median	332,8	1798,4	67,1
Female (n=10)	Average	379,11	1973,35	70,22
	Standard deviation	63,48631436	194,1718841	14,45358856
	Median	368,65	1989,1	70,55

schub /s x ʏ p/ ‘scale’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	366,35	1663,2	73,65
	Standard deviation	15,06137444	220,334473	6,858935778
	Median	366,35	1663,2	73,65
Female (n=10)	Average	406,49	1955,68	76,07
	Standard deviation	48,76716678	180,4830174	8,852124666
	Median	400,65	2017,1	74,2

kus /k ʏ s/ ‘kiss’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	330,65	1895,5	71,65
	Standard deviation	40,94148263	230,0925466	3,323401872
	Median	330,65	1895,5	71,65
Female (n=10)	Average	387,08	2005,76	88,31
	Standard deviation	54,11700903	176,3075734	18,5511575
	Median	384,35	1969,7	82,55

brug /b r ʏ x/‘bridge’		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	362,5	1731,65	73,65
	Standard deviation	49,7803174	126,9256672	7,424621202
	Median	362,5	1731,65	73,65
Female (n=10)	Average	406,28	2002,06	84,5
	Standard deviation	46,99531419	254,2523121	14,12011961
	Median	406,05	2017,55	85,15

2 English production task

2.1 Ghent participants

key /k iː/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	268,65	2316,775	172,775
	Standard deviation	40,51308431	55,3823904	30,20478715
	Median	261,55	2320,05	179,25
Female (n=8)	Average	304,05	2744,4	172,75
	Standard deviation	44,91270899	160,3607807	15,07751401
	Median	300,55	2718,4	174,35

sheep /ʃ iː p/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	286,7	2310,575	135,8
	Standard deviation	45,07467877	95,85768531	8,888569439
	Median	269,45	2319,95	135,55
Female (n=8)	Average	355,8625	2653,4375	126,5
	Standard deviation	32,60516775	173,9385519	13,53028349
	Median	357,85	2609,5	130,2

peace /p iː s/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	283,725	2328,6	155,925
	Standard deviation	32,39221872	200,9425623	12,3243323
	Median	276,85	2367,65	159,9
Female (n=8)	Average	310,7875	2563,425	151,0875
	Standard deviation	38,60989649	191,8868994	18,07116389
	Median	300,65	2616,45	150,4

leaf /l iː f/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	303,775	2296,125	148,35
	Standard deviation	30,82600363	53,84096179	24,32591211
	Median	305,4	2294,8	147,35
Female (n=8)	Average	322,075	2684,2375	158,6375
	Standard deviation	31,22676279	107,0709503	24,85184313
	Median	323,55	2679,75	164,6

spit /s p ɪ t/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	420,375	1809,075	70,5
	Standard deviation	17,05801376	138,9355336	12,49186402
	Median	417,9	1775,15	66,95
Female (n=8)	Average	464,7	2047,0875	75,375
	Standard deviation	49,00848323	112,2935367	9,351814186
	Median	464,95	2017,1	75,95

fist /f ɪ s t/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	420,575	1857,175	77,725
	Standard deviation	31,24935999	200,0957	11,61991824
	Median	427,9	1781,35	80,7
Female (n=8)	Average	482,0875	2027,6	84,25
	Standard deviation	52,53954803	125,1105226	14,79150335
	Median	492,65	2046,95	83,6

six /s ɪ k s/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	406	2111,775	72,625
	Standard deviation	28,39917839	71,62771228	15,43640178
	Median	411,15	2089,9	75,95
Female (n=8)	Average	443,1875	2375,15	83,775
	Standard deviation	60,76802349	133,5139907	13,09086377
	Median	455,7	2399,9	82,95

pig /p ɪ g/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	391,85	2098,275	115,425
	Standard deviation	36,04816223	56,42590274	20,93503841
	Median	386,65	2089,6	109
Female (n=8)	Average	480,1375	2349,675	112,475
	Standard deviation	32,29077433	160,1813236	21,3083317
	Median	477,5	2393,7	105,85

book /b ʊ k/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	413,025	1039,9	68,625
	Standard deviation	38,24233039	213,474979	10,22395064
	Median	414,15	1060,65	66
Female (n=8)	Average	447,375	1152,9375	91,9875
	Standard deviation	61,72641828	251,4301263	20,69109108
	Median	434,5	1047,8	92,05

foot /f ʊ t/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	378,15	1285,15	72,425
	Standard deviation	29,33331439	305,5539396	10,20273656
	Median	376,95	1318,3	72,05
Female (n=8)	Average	419,2875	1369,425	98,4375
	Standard deviation	38,81061848	249,8695703	26,73541993
	Median	413,45	1264,4	88,8

hook /h ʊ k/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	392,075	1078,725	78,45
	Standard deviation	55,33855648	232,2745125	13,09592812
	Median	382,15	1102,95	76,5
Female (n=8)	Average	434,0625	1169,975	91,3875
	Standard deviation	41,40275828	167,6991417	12,63175562
	Median	442,2	1116,85	95,8

cook /k ʊ k/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	365,1	1070,2	74,675
	Standard deviation	9,627391478	240,665217	12,40520724
	Median	366,4	1001,6	70,55
Female (n=8)	Average	413,9875	1044,375	89,2625
	Standard deviation	53,13603667	101,0263014	16,89876897
	Median	425,3	1065,35	90,05

hug /h ʌ g/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	570,375	1268,875	116,325
	Standard deviation	15,74047331	66,73761433	34,54662309
	Median	565,8	1293,75	116,4
Female (n=8)	Average	710,075	1512,8375	118,8375
	Standard deviation	143,1691881	167,4676168	23,27002594
	Median	720,9	1509,75	113,1

duck /d ʌ k/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	624,975	1289,125	99,45
	Standard deviation	44,82977991	77,71170547	10,29579202
	Median	625,75	1254,9	97,65
Female (n=8)	Average	649,075	1582,375	107,975
	Standard deviation	110,7638769	110,4020671	19,8043105
	Median	668,3	1548,35	106,85

cup /k ʌ p/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	644,325	1200,55	87,3
	Standard deviation	57,20631521	41,86848457	12,9645671
	Median	623,45	1211,1	90,2
Female (n=8)	Average	670,775	1471,1875	98,55
	Standard deviation	88,92254655	149,2545614	13,29435971
	Median	679,05	1460,55	94,95

bus /b ʌ s/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	567,35	1175,4	102,325
	Standard deviation	97,0513438	161,039519	20,50144304
	Median	556,25	1124,85	100,35
Female (n=8)	Average	710,2125	1475,4875	106,7125
	Standard deviation	68,58297243	113,9781989	17,11752965
	Median	706,7	1463,75	108,55

2.2. Antwerp participants

key /k iː/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	323,7	2421,35	162,6
	Standard deviation	59,96265504	244,8710783	22,76883835
	Median	323,7	2421,35	162,6
Female (n=10)	Average	365,2	2706,16	189,54
	Standard deviation	37,13037092	144,053649	24,09726587
	Median	356,95	2674,55	186,6

sheep /ʃ iː p/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	325,7	2458,2	92,9
	Standard deviation	12,44507935	226,9812768	7,636753237
	Median	325,7	2458,2	92,9
Female (n=10)	Average	329,39	2562,4	131,91
	Standard deviation	44,28278948	223,4117922	28,11130457
	Median	320,2	2607,25	123,4

peace /p iː s/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	309,9	2430,9	145,15
	Standard deviation	24,8901587	203,3639103	28,92066735
	Median	309,9	2430,9	145,15
Female (n=10)	Average	343,24	2590,54	155,33
	Standard deviation	37,82216281	171,9137781	14,20321169
	Median	348,3	2576,15	155,45

leaf /l iː f/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	309,35	2358,9	127,55
	Standard deviation	38,82016229	239,5677775	0,494974747
	Median	309,35	2358,9	127,55
Female (n=10)	Average	337,95	2690,46	163,66
	Standard deviation	59,88112762	132,5723299	23,83919275
	Median	328,7	2682,8	157,5

spit /s p ɪ t/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	390,7	1933	67,35
	Standard deviation	57,98275606	254,1341772	17,32411614
	Median	390,7	1933	67,35
Female (n=10)	Average	436,08	2249,82	68,26
	Standard deviation	34,96518268	126,5971283	9,459410365
	Median	448,5	2213,3	70,25

fist /f ɪ s t/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	372,75	2018,5	66,5
	Standard deviation	7,283199846	412,3846748	3,676955262
	Median	372,75	2018,5	66,5
Female (n=10)	Average	431,83	2313,28	76,83
	Standard deviation	58,45463673	123,1931257	8,761157965
	Median	414,75	2336,4	78,45

six /s ɪ k s/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	395	2140,15	74,6
	Standard deviation	6,505382387	368,190501	16,68772004
	Median	395	2140,15	74,6
Female (n=10)	Average	442,96	2423,91	78,68
	Standard deviation	49,3622933	109,7499729	8,312346908
	Median	454,7	2385,35	78,5

pig /p ɪ g/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	355,45	2294,65	84,85
	Standard deviation	20,01112191	266,2257031	7,424621202
	Median	355,45	2294,65	84,85
Female (n=10)	Average	439,35	2475,95	95,9
	Standard deviation	44,63153469	118,210173	9,909479188
	Median	454,3	2513,9	95,5

book /b ʊ k/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	390,65	1038,45	68,2
	Standard deviation	32,59762261	54,9421969	7,778174593
	Median	390,65	1038,45	68,2
Female (n=10)	Average	426,82	1013,71	99,34
	Standard deviation	87,88396137	89,09062115	14,54504116
	Median	446,25	1028,55	95,25

foot /f ʊ t/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	408,15	1195,45	77,95
	Standard deviation	21,14249276	83,93357493	10,6773124
	Median	408,15	1195,45	77,95
Female (n=10)	Average	443,93	1221,79	91,4
	Standard deviation	49,62906966	226,4630774	11,28548724
	Median	442,05	1206,05	87,25

hook /h ʊ k/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	393,95	1088,8	73,55
	Standard deviation	12,51579003	96,3079436	1,626345597
	Median	393,95	1088,8	73,55
Female (n=10)	Average	451,3	1112,83	91,4
	Standard deviation	31,18165415	241,1109106	18,41985644
	Median	449,65	1078,65	93,1

cook /k ʊ k/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	372,7	1018,45	58,95
	Standard deviation	18,52619767	33,30472939	14,07142495
	Median	372,7	1018,45	58,95
Female (n=10)	Average	477,29	1145,31	87,18
	Standard deviation	34,99804598	122,9380192	8,598035951
	Median	476,2	1108,75	88,8

hug /h ʌ g/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	500,8	1328,4	103,45
	Standard deviation	19,65756852	83,29717882	6,576093065
	Median	500,8	1328,4	103,45
Female (n=10)	Average	647,12	1627,04	112,14
	Standard deviation	66,26164971	96,62691367	12,13985173
	Median	651,2	1605,75	111,1

duck /d ʌ k/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	554,1	1424,5	96,25
	Standard deviation	32,52691193	100,6920056	0,636396103
	Median	554,1	1424,5	96,25
Female (n=10)	Average	631,24	1616,37	118,75
	Standard deviation	109,2072464	227,987266	29,41470721
	Median	627,4	1640,15	113,2

cup /k ʌ p/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	574,35	1319,6	69,3
	Standard deviation	33,87041482	196,5756852	14,99066376
	Median	574,35	1319,6	69,3
Female (n=10)	Average	606,33	1529,16	93,78
	Standard deviation	125,5109119	124,2337063	9,158820885
	Median	584,7	1523,9	91,3

bus /b ʌ s/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	556,05	1278	98,1
	Standard deviation	61,87184335	76,22611101	2,545584412
	Median	556,05	1278	98,1
Female (n=10)	Average	620,31	1584,55	117,13
	Standard deviation	75,08782192	135,0262958	19,34143796
	Median	633,2	1583,05	111,45

Appendix H. Production experiment: average values (vowels)

1 Dutch production task

1.1 Ghent participants

/i/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	289,4625	2178,9625	75,16875
	Standard deviation	38,5265601	142,9000017	9,16818548
	Median	291,8	2147,35	76,8
Female (n=8)	Average	315,53125	2598,853125	83,425
	Standard deviation	38,37606308	132,5993115	13,02161974
	Median	298,65	2595,35	82,4

/ɪ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	424,1375	1826,575	71,5
	Standard deviation	26,73920655	127,0634986	12,50050666
	Median	423,35	1789,2	67,35
Female (n=8)	Average	480,475	2115,178125	72,4625
	Standard deviation	36,80199854	128,3771328	12,44162498
	Median	481,15	2086,1	75,6

/y/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	282,2625	1983,74375	165,74375
	Standard deviation	41,65763435	138,2126717	22,79809842
	Median	275,5	1953,15	163,45
Female (n=8)	Average	316,70625	2138,5375	171,46875
	Standard deviation	29,18308102	203,1246538	26,30259937
	Median	318,65	2110,9	175,7

/ʏ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	432,8125	1553,7125	79,48125
	Standard deviation	49,94187121	106,3977561	10,84209197
	Median	418,8	1558,3	76,55
Female (n=8)	Average	445,50625	1777,946875	83,58125
	Standard deviation	40,66449619	125,410567	14,83455942
	Median	438,85	1781	85,6

1.2 Antwerp participants

/i/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	309,25	2439,225	124,025
	Standard deviation	38,20620742	173,3500155	9,250752865
	Median	296,5	2433,8	122,05
Female (n=10)	Average	327,1475	2618,1625	140,125641
	Standard deviation	43,92989163	161,0691438	27,16061117
	Median	316,1	2646,5	135,9

/ɪ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	346,2	2211,5125	67,2625
	Standard deviation	23,59933413	155,6538599	10,26143091
	Median	350,25	2248,6	63,7
Female (n=10)	Average	399,395	2449,5925	69,15
	Standard deviation	59,57134034	182,39316	10,85435042
	Median	408	2442,75	68,95

/y/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	313,425	1960,2625	157,625
	Standard deviation	32,08064615	110,4639688	16,57327884
	Median	311,7	1942,55	159,35
Female (n=10)	Average	339,0225	2133,695	180,0175
	Standard deviation	59,48041474	190,326746	27,2593422
	Median	329,65	2082,35	177,75

/ʏ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	348,075	1772,1875	71,5125
	Standard deviation	35,39627704	158,8428823	5,612088356
	Median	357,65	1798,4	70,7
Female (n=10)	Average	394,74	1984,2125	79,775
	Standard deviation	53,02531954	196,8468234	15,59016373
	Median	394,95	1989,1	78,5

2 English production task

1.1 Ghent participants

/iː/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	285,7125	2313,01875	153,2125
	Standard deviation	36,06183347	106,0789091	23,19838716
	Median	278,95	2299,25	147,35
Female (n=8)	Average	323,19375	2661,371875	152,24375
	Standard deviation	40,77228611	167,1719865	24,41957091
	Median	317,8	2683,35	151,4

/ɪ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	409,7	1969,075	84,06875
	Standard deviation	28,72696759	183,2443742	23,46201806
	Median	412,55	2048,65	81,7
Female (n=8)	Average	467,528125	2199,878125	88,96875
	Standard deviation	49,85633059	208,8137119	20,33131126
	Median	468,55	2169,5	87,6

/ʊ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	387,0875	1118,49375	73,54375
	Standard deviation	37,83976524	245,4761916	10,97038248
	Median	374,5	1072,75	72,05
Female (n=8)	Average	428,678125	1184,178125	92,76875
	Standard deviation	49,00256566	226,3440697	19,2536622
	Median	436,55	1141,75	91,6

/ʌ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=4)	Average	601,75625	1233,4875	101,35
	Standard deviation	64,76939574	99,92214219	22,16369404
	Median	593,85	1239,85	95,15
Female (n=8)	Average	685,034375	1510,471875	108,01875
	Standard deviation	104,6803571	138,1976617	19,25337737
	Median	692,35	1528,1	107,5

1.2 Antwerp participants

/iː/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	317,1625	2417,3375	132,05
	Standard deviation	30,08497192	177,6036513	31,00497656
	Median	322,2	2413	127,55
Female (n=10)	Average	343,945	2637,39	160,11
	Standard deviation	45,91063197	176,4164708	30,48194758
	Median	348,15	2650,1	156,25

/ɪ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	379,7428571	2096,575	73,325
	Standard deviation	30,92651813	290,0875767	12,42529103
	Median	377,9	2109,55	74,35
Female (n=10)	Average	437,555	2365,74	79,9175
	Standard deviation	45,93388197	146,0763443	13,42194696
	Median	448,5	2383,3	78,65

/ʊ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	391,3625	1085,2875	69,6625
	Standard deviation	21,66003149	91,07779151	10,53252615
	Median	389,5	1059,65	71,4
Female (n=10)	Average	449,835	1123,41	92,33
	Standard deviation	56,5365711	190,4973676	13,90735279
	Median	451,45	1087,35	91,15

/ʌ/		F1 (Hz)	F2 (Hz)	Duration (ms)
Male (n=2)	Average	546,325	1337,625	91,775
	Standard deviation	42,16575286	109,9055016	15,48167119
	Median	540,75	1342,6	96,5
Female (n=10)	Average	626,25	1589,28	110,45
	Standard deviation	94,50240263	152,9868726	20,98702896
	Median	638,3	1576	107,7

Appendix I. Production Experiment: SPSS outputs (per vowel)

1. Dutch production task

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
iVowelV.F1	1,00	40	327,1475	43,92989	6,94593
	2,00	32	315,5312	38,37606	6,78399
iVowelV.F2	1,00	40	2618,1625	161,06914	25,46727
	2,00	32	2598,8531	132,59931	23,44047
iVowelV.L	1,00	40	138,5100	28,69138	4,53651
	2,00	32	83,4250	13,02162	2,30192

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
iVowelV.F1	Equal variances assumed	,040	,841	1,178	70	,243	11,61625	9,85729	-8,04350	31,27600
	Equal variances not assumed			1,196	69,421	,236	11,61625	9,70919	-7,75097	30,98347
iVowelV.F2	Equal variances assumed	1,613	,208	,546	70	,587	19,30938	35,37002	-51,23390	89,85265
	Equal variances not assumed			,558	69,929	,579	19,30938	34,61268	-49,72465	88,34340
iVowelV.L	Equal variances assumed	17,035	,000	10,053	70	,000	55,08500	5,47926	44,15696	66,01304
	Equal variances not assumed			10,828	56,921	,000	55,08500	5,08711	44,89793	65,27207

SPSS output for Dutch /i/ ('iVowel') produced by the Antwerp (WoonplaatsV or 'City of residence'= 1) and Ghent (WoonplaatsV or 'City of residence'= 2) female participants ('V' for 'vrouwelijk', Dutch for 'female') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
iVowelM.F1	1,00	8	309,2500	38,20621	13,50793
	2,00	16	289,4625	38,52656	9,63164
iVowelM.F2	1,00	8	2439,2250	173,35002	61,28849
	2,00	16	2178,9625	142,90000	35,72500
iVowelM.L	1,00	8	124,0250	9,25075	3,27064
	2,00	16	75,1688	9,16819	2,29205

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
iVowelM.F1	Equal variances assumed	,001	,976	1,189	22	,247	19,78750	16,63848	-14,71859	54,29359
	Equal variances not assumed			1,193	14,213	,253	19,78750	16,59014	-15,74489	55,31989
iVowelM.F2	Equal variances assumed	1,278	,270	3,922	22	,001	260,26250	66,35762	122,64523	397,87977
	Equal variances not assumed			3,669	11,923	,003	260,26250	70,94050	105,58506	414,93994
iVowelM.L	Equal variances assumed	,057	,814	12,271	22	,000	48,85625	3,98135	40,59943	57,11307
	Equal variances not assumed			12,233	13,989	,000	48,85625	3,99381	40,28976	57,42274

SPSS output for Dutch /i/ ('iVowel') produced by the Antwerp (WoonplaatsM or 'City of residence'= 1) and Ghent (WoonplaatsM or 'City of residence'= 2) male participants ('M') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
IVowelV.F1	1,00	40	399,3950	59,57134	9,41906
	2,00	32	480,4750	36,80200	6,50574
IVowelV.F2	1,00	40	2449,5925	182,39316	28,83889
	2,00	32	2115,1781	128,37713	22,69409
IVowelV.L	1,00	40	69,1500	10,85435	1,71622
	2,00	32	72,4625	12,44162	2,19939

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IVowelV.F1	Equal variances assumed	11,080	,001	-6,734	70	,000	-81,08000	12,03968	-105,09238	-57,06762
	Equal variances not assumed			-7,083	66,148	,000	-81,08000	11,44741	-103,93454	-58,22546
IVowelV.F2	Equal variances assumed	,861	,357	8,773	70	,000	334,41437	38,11979	258,38686	410,44189
	Equal variances not assumed			9,113	68,979	,000	334,41437	36,69745	261,20456	407,62419
IVowelV.L	Equal variances assumed	1,928	,169	-1,206	70	,232	-3,31250	2,74742	-8,79206	2,16706
	Equal variances not assumed			-1,187	61,979	,240	-3,31250	2,78976	-8,88918	2,26418

SPSS output for Dutch /ɪ/ ('IVowel') produced by the Antwerp (WoonplaatsV or 'City of residence'= 1) and Ghent (WoonplaatsV or 'City of residence'= 2) female participants ('V' for 'vrouwelijk', Dutch for 'female') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
IVowelM.F1	1,00	8	346,2000	23,59933	8,34362
	2,00	16	424,1375	26,73921	6,68480
IVowelM.F2	1,00	8	2211,5125	155,65386	55,03195
	2,00	16	1826,5750	127,06350	31,76587
IVowelM.L	1,00	8	67,2625	10,26143	3,62796
	2,00	16	71,5000	12,50051	3,12513

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IVowelM.F1	Equal variances assumed	,189	,668	-6,981	22	,000	-77,93750	11,16379	-101,08978	-54,78522
	Equal variances not assumed			-7,290	15,827	,000	-77,93750	10,69124	-100,62201	-55,25299
IVowelM.F2	Equal variances assumed	,649	,429	6,498	22	,000	384,93750	59,24050	262,08023	507,79477
	Equal variances not assumed			6,058	11,829	,000	384,93750	63,54200	246,26899	523,60601
IVowelM.L	Equal variances assumed	1,598	,219	-,827	22	,417	-4,23750	5,12432	-14,86469	6,38969
	Equal variances not assumed			-,885	16,900	,389	-4,23750	4,78838	-14,34465	5,86965

SPSS output for Dutch /ɪ/ ('IVowel') produced by the Antwerp (WoonplaatsM or 'City of residence'= 1) and Ghent (WoonplaatsM or 'City of residence'= 2) male participants ('M') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
yVowelV.F1	1,00	40	339,0225	59,48041	9,40468
	2,00	32	316,7063	29,18308	5,15889
yVowelV.F2	1,00	40	2133,6950	190,32675	30,09330
	2,00	32	2138,5375	203,12465	35,90771
yVowelV.L	1,00	40	180,0175	27,25934	4,31008
	2,00	32	171,4688	26,30260	4,64969

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
yVowelV.F1	Equal variances assumed	9,498	,003	1,942	70	,056	22,31625	11,49309	-,60600	45,23850
	Equal variances not assumed			2,080	59,252	,042	22,31625	10,72670	,85408	43,77842
yVowelV.F2	Equal variances assumed	,030	,864	-,104	70	,917	-4,84250	46,50860	-97,60097	87,91597
	Equal variances not assumed			-,103	64,534	,918	-4,84250	46,85051	-98,42223	88,73723
yVowelV.L	Equal variances assumed	,068	,796	1,343	70	,184	8,54875	6,36563	-4,14710	21,24460
	Equal variances not assumed			1,348	67,530	,182	8,54875	6,34006	-4,10424	21,20174

SPSS output for Dutch /y/ ('yVowel') produced by the Antwerp (WoonplaatsV or 'City of residence'= 1) and Ghent (WoonplaatsV or 'City of residence'= 2) female participants ('V' for 'vrouwelijk', Dutch for 'female') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
yVowelM.F1	1,00	8	313,4250	32,08065	11,34222
	2,00	16	282,2625	41,65763	10,41441
yVowelM.F2	1,00	8	1960,2625	110,46397	39,05491
	2,00	16	1983,7438	138,21267	34,55317
yVowelM.L	1,00	8	157,6250	16,57328	5,85954
	2,00	16	165,7438	22,79810	5,69952

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
yVowelM.F1	Equal variances assumed	,838	,370	1,852	22	,078	31,16250	16,83000	-3,74079	66,06579
	Equal variances not assumed			2,024	17,856	,058	31,16250	15,39824	-1,20673	63,53173
yVowelM.F2	Equal variances assumed	1,145	,296	-,417	22	,681	-23,48125	56,30355	-140,24766	93,28516
	Equal variances not assumed			-,450	17,301	,658	-23,48125	52,14602	-133,35425	86,39175
yVowelM.L	Equal variances assumed	1,127	,300	-,892	22	,382	-8,11875	9,10124	-26,99356	10,75606
	Equal variances not assumed			-,993	18,700	,333	-8,11875	8,17428	-25,24629	9,00879

SPSS output for Dutch /y/ ('yVowel') produced by the Antwerp (WoonplaatsM or 'City of residence'= 1) and Ghent (WoonplaatsM or 'City of residence'= 2) male participants ('M') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
YVowelV.F1	1,00	40	394,7400	53,02532	8,38404
	2,00	32	445,5062	40,66450	7,18854
YVowelV.F2	1,00	40	1984,2125	196,84682	31,12422
	2,00	32	1777,9469	125,41057	22,16967
YVowelV.L	1,00	40	79,7750	15,59016	2,46502
	2,00	32	83,5813	14,83456	2,62240

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
YVowelV.F1	Equal variances assumed	2,171	,145	-4,464	70	,000	-50,76625	11,37139	-73,44578	-28,08672
	Equal variances not assumed			-4,597	69,896	,000	-50,76625	11,04387	-72,79314	-28,73936
YVowelV.F2	Equal variances assumed	5,966	,017	5,147	70	,000	206,26563	40,07676	126,33505	286,19620
	Equal variances not assumed			5,398	66,936	,000	206,26563	38,21271	129,99141	282,53984
YVowelV.L	Equal variances assumed	,044	,835	-1,052	70	,297	-3,80625	3,61926	-11,02464	3,41214
	Equal variances not assumed			-1,058	67,868	,294	-3,80625	3,59907	-10,98834	3,37584

SPSS output for Dutch /ʏ/ ('YVowel') produced by the Antwerp (WoonplaatsV or 'City of residence'= 1) and Ghent (WoonplaatsV or 'City of residence'= 2) female participants ('V' for 'vrouwelijk', Dutch for 'female') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
YVowelM.F1	1,00	8	348,0750	35,39628	12,51447
	2,00	16	432,8125	49,94187	12,48547
YVowelM.F2	1,00	8	1772,1875	158,84288	56,15944
	2,00	16	1553,7125	106,39776	26,59944
YVowelM.L	1,00	8	71,5125	5,61209	1,98417
	2,00	16	79,4813	10,84209	2,71052

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
YVowelM.F1	Equal variances assumed	,349	,561	-4,271	22	,000	-84,73750	19,83951	-125,88213	-43,59287
	Equal variances not assumed			-4,793	19,059	,000	-84,73750	17,67764	-121,72952	-47,74548
YVowelM.F2	Equal variances assumed	,679	,419	4,021	22	,001	218,47500	54,33676	105,78747	331,16253
	Equal variances not assumed			3,516	10,252	,005	218,47500	62,14027	80,47817	356,47183
YVowelM.L	Equal variances assumed	2,104	,161	-1,938	22	,066	-7,96875	4,11179	-16,49608	,55858
	Equal variances not assumed			-2,372	21,905	,027	-7,96875	3,35915	-14,93695	-1,00055

SPSS output for Dutch /ʏ/ ('YVowel') produced by the Antwerp (WoonplaatsM or 'City of residence'= 1) and Ghent (WoonplaatsM or 'City of residence'= 2) male participants ('M') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

2. English production task

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
iLongVowelV.F1	1,00	40	343,9450	45,91063	7,25911
	2,00	32	323,1938	40,77229	7,20759
iLongVowelV.F2	1,00	40	2637,3900	176,41647	27,89389
	2,00	32	2661,3719	167,17199	29,55211
iLongVowelV.L	1,00	40	160,1100	30,48195	4,81962
	2,00	32	152,2437	24,41957	4,31681

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
iLongVowelV.F1	Equal variances assumed	,217	,642	2,002	70	,049	20,75125	10,36666	,07561	41,42689
	Equal variances not assumed			2,029	69,195	,046	20,75125	10,22957	,34485	41,15765
iLongVowelV.F2	Equal variances assumed	,022	,882	-,587	70	,559	-23,98188	40,88437	-105,52319	57,55944
	Equal variances not assumed			-,590	67,963	,557	-23,98188	40,63738	-105,07329	57,10954
iLongVowelV.L	Equal variances assumed	,778	,381	1,186	70	,240	7,86625	6,63125	-5,35937	21,09187
	Equal variances not assumed			1,216	69,999	,228	7,86625	6,47021	-5,03818	20,77068

SPSS output for English /iː/ ('iLongVowel') produced by the Antwerp (WoonplaatsV or 'City of residence'= 1) and Ghent (WoonplaatsV or 'City of residence'= 2) female participants ('V' for 'vrouwelijk', Dutch for 'female') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
iLongVowelM.F1	1,00	8	317,1625	30,08497	10,63664
	2,00	16	285,7125	36,06183	9,01546
iLongVowelM.F2	1,00	8	2417,3375	177,60365	62,79237
	2,00	16	2313,0188	106,07891	26,51973
iLongVowelM.L	1,00	8	132,0500	31,00498	10,96191
	2,00	16	153,2125	23,19839	5,79960

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
iLongVowelM.F1	Equal variances assumed	,546	,468	2,119	22	,046	31,45000	14,84080	,67207	62,22793
	Equal variances not assumed			2,256	16,658	,038	31,45000	13,94334	1,98608	60,91392
iLongVowelM.F2	Equal variances assumed	9,941	,005	1,810	22	,084	104,31875	57,62283	-15,18368	223,82118
	Equal variances not assumed			1,530	9,578	,158	104,31875	68,16288	-48,47022	257,10772
iLongVowelM.L	Equal variances assumed	,607	,444	-1,884	22	,073	-21,16250	11,23167	-44,45555	2,13055
	Equal variances not assumed			-1,706	11,063	,116	-21,16250	12,40157	-48,43929	6,11429

SPSS output for English /i:/ ('iLongVowel') produced by the Antwerp (WoonplaatsM or 'City of residence'= 1) and Ghent (WoonplaatsM or 'City of residence'= 2) male participants ('M') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
IVowelEV.F1	1,00	40	437,5550	45,93388	7,26278
	2,00	32	467,5281	49,85633	8,81344
IVowelEV.F2	1,00	40	2365,7400	146,07634	23,09670
	2,00	32	2199,8781	208,81371	36,91340
IVowelEV.L	1,00	40	79,9175	13,42195	2,12220
	2,00	32	88,9688	20,33131	3,59410

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IVowelEV.F1	Equal variances assumed	,001	,982	-2,649	70	,010	-29,97313	11,31560	-52,54138	-7,40487
	Equal variances not assumed			-2,625	63,955	,011	-29,97313	11,42036	-52,78823	-7,15802
IVowelEV.F2	Equal variances assumed	7,464	,008	3,959	70	,000	165,86188	41,89168	82,31155	249,41220
	Equal variances not assumed			3,809	53,506	,000	165,86188	43,54373	78,54348	253,18027
IVowelEV.L	Equal variances assumed	3,975	,050	-2,267	70	,026	-9,05125	3,99285	-17,01475	-1,08775
	Equal variances not assumed			-2,169	51,417	,035	-9,05125	4,17388	-17,42902	-,67348

SPSS output for English /ɪ/ ('IVowel') produced by the Antwerp (WoonplaatsV or 'City of residence'= 1) and Ghent (WoonplaatsV or 'City of residence'= 2) female participants ('V' for 'vrouwelijk', Dutch for 'female') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
IVowelEM.F1	1,00	8	332,2750	137,27853	48,53529
	2,00	16	409,7000	28,72697	7,18174
IVowelEM.F2	1,00	8	2096,5750	290,08758	102,56145
	2,00	16	1969,0750	183,24437	45,81109
IVowelEM.L	1,00	8	73,3250	12,42529	4,39300
	2,00	16	84,0688	23,46202	5,86550

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
IVowelEM.F1	Equal variances assumed	5,270	,032	-2,208	22	,038	-77,42500	35,06848	-150,15258	-4,69742
	Equal variances not assumed			-1,578	7,308	,157	-77,42500	49,06375	-192,45768	37,60768
IVowelEM.F2	Equal variances assumed	2,410	,135	1,321	22	,200	127,50000	96,50423	-72,63752	327,63752
	Equal variances not assumed			1,135	9,888	,283	127,50000	112,32767	-123,16593	378,16593
IVowelEM.L	Equal variances assumed	1,431	,244	-1,204	22	,241	-10,74375	8,92092	-29,24460	7,75710
	Equal variances not assumed			-1,466	21,829	,157	-10,74375	7,32821	-25,94841	4,46091

SPSS output for English /ɪ/ ('IVowel') produced by the Antwerp (WoonplaatsM or 'City of residence'= 1) and Ghent (WoonplaatsM or 'City of residence'= 2) male participants ('M') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
UVowelV.F1	1,00	40	449,8350	56,53657	8,93922
	2,00	32	428,6781	49,00257	8,66251
UVowelV.F2	1,00	40	1123,4100	190,49737	30,12028
	2,00	32	1184,1781	226,34407	40,01236
UVowelV.L	1,00	40	92,3300	13,90735	2,19895
	2,00	32	92,7688	19,25366	3,40360

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
UVowelV.F1	Equal variances assumed	,044	,834	1,673	70	,099	21,15687	12,64869	-4,07014	46,38389
	Equal variances not assumed			1,700	69,516	,094	21,15687	12,44784	-3,67259	45,98634
UVowelV.F2	Equal variances assumed	,649	,423	-1,237	70	,220	-60,76812	49,12734	-158,74952	37,21327
	Equal variances not assumed			-1,213	60,616	,230	-60,76812	50,08213	-160,92636	39,39011
UVowelV.L	Equal variances assumed	3,684	,059	-,112	70	,911	-,43875	3,91101	-8,23900	7,36150
	Equal variances not assumed			-,108	54,704	,914	-,43875	4,05214	-8,56041	7,68291

SPSS output for English /ʊ/ ('UVowel') produced by the Antwerp (WoonplaatsV or 'City of residence'= 1) and Ghent (WoonplaatsV or 'City of residence'= 2) female participants ('V' for 'vrouwelijk', Dutch for 'female') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
UVowelM.F1	1,00	8	391,3625	21,66003	7,65798
	2,00	16	387,0875	37,83977	9,45994
UVowelM.F2	1,00	8	1085,2875	91,07779	32,20086
	2,00	16	1118,4938	245,47619	61,36905
UVowelM.L	1,00	8	69,6625	10,53253	3,72381
	2,00	16	73,5437	10,97038	2,74260

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
UVowelM.F1	Equal variances assumed	2,732	,113	,294	22	,771	4,27500	14,52715	-25,85247	34,40247
	Equal variances not assumed			,351	21,404	,729	4,27500	12,17108	-21,00706	29,55706
UVowelM.F2	Equal variances assumed	8,210	,009	-,367	22	,717	-33,20625	90,54497	-220,98502	154,57252
	Equal variances not assumed			-,479	20,988	,637	-33,20625	69,30408	-177,33717	110,92467
UVowelM.L	Equal variances assumed	,323	,575	-,827	22	,417	-3,88125	4,69082	-13,60941	5,84691
	Equal variances not assumed			-,839	14,643	,415	-3,88125	4,62478	-13,75970	5,99720

SPSS output for English /ʊ/ ('UVowel') produced by the Antwerp (WoonplaatsM or 'City of residence'= 1) and Ghent (WoonplaatsM or 'City of residence'= 2) male participants ('M') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsV	N	Mean	Std. Deviation	Std. Error Mean
BusVowelV.F1	1,00	40	626,2500	94,50240	14,94214
	2,00	32	685,0344	104,68036	18,50505
BusVowelV.F2	1,00	40	1589,2800	152,98687	24,18935
	2,00	32	1510,4719	138,19766	24,43013
BusVowelV.L	1,00	40	110,4500	20,98703	3,31834
	2,00	32	108,0187	19,25338	3,40355

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
BusVowelV.F1	Equal variances assumed	,001	,980	-2,500	70	,015	-58,78438	23,51282	-105,67923	-11,88952
	Equal variances not assumed			-2,472	63,235	,016	-58,78438	23,78454	-106,31053	-11,25822
BusVowelV.F2	Equal variances assumed	,010	,922	2,266	70	,027	78,80812	34,77434	9,45288	148,16337
	Equal variances not assumed			2,292	68,923	,025	78,80812	34,37958	10,22137	147,39488
BusVowelV.L	Equal variances assumed	,079	,779	,507	70	,614	2,43125	4,79977	-7,14159	12,00409
	Equal variances not assumed			,511	68,644	,611	2,43125	4,75348	-7,05255	11,91505

SPSS output for English /ʌ/ ('BusVowel') produced by the Antwerp (WoonplaatsV or 'City of residence'= 1) and Ghent (WoonplaatsV or 'City of residence'= 2) female participants ('V' for 'vrouwelijk', Dutch for 'female') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Group Statistics

	WoonplaatsM	N	Mean	Std. Deviation	Std. Error Mean
BusVowelM.F1	1,00	8	546,3250	42,16575	14,90784
	2,00	16	601,7562	64,76940	16,19235
BusVowelM.F2	1,00	8	1337,6250	109,90550	38,85746
	2,00	16	1233,4875	99,92214	24,98054
BusVowelM.L	1,00	8	91,7750	15,48167	5,47360
	2,00	16	101,3500	22,16369	5,54092

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
BusVowelM.F1	Equal variances assumed	,700	,412	-2,187	22	,040	-55,43125	25,34508	-107,99373	-2,86877
	Equal variances not assumed			-2,518	20,163	,020	-55,43125	22,00991	-101,31932	-9,54318
BusVowelM.F2	Equal variances assumed	,187	,670	2,330	22	,029	104,13750	44,68841	11,45940	196,81560
	Equal variances not assumed			2,254	12,950	,042	104,13750	46,19448	4,30083	203,97417
BusVowelM.L	Equal variances assumed	1,275	,271	-1,090	22	,287	-9,57500	8,78057	-27,78479	8,63479
	Equal variances not assumed			-1,229	19,259	,234	-9,57500	7,78859	-25,86186	6,71186

SPSS output for English /ʌ/ ('BusVowel') produced by the Antwerp (WoonplaatsM or 'City of residence'= 1) and Ghent (WoonplaatsM or 'City of residence'= 2) male participants ('M') with regard to the F1, F2, and duration ('L' for 'length') features. The appropriate t statistic (t), degrees of freedom (df) and p-value (Sig. (2-tailed)) - depending on the outcome of Levene's test for equality of variances - are indicated with a green box.

Appendix J. Production experiment: p-values (stimuli)

Since presenting screenshots from the SPSS outputs for all stimuli would occupy too much space, this appendix is limited to the presentation of the p-values of the t-tests for equality of means, conducted for the F1, F2, and duration ('L' for 'length') features per stimulus. Results that are significant (on the 5%-significance level, so $p \leq 0.050$) are indicated in green. n^1 refers the number of Antwerp participants in the sample, n^2 refers to the number of Ghent participants in the sample.

1 Dutch production task

kies/k i s/ 'molar'	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.326	0.088	0.006
Female (n^1=10; n^2=8)	0.572	0.691	< 0.001

brief /b r i f/'letter'	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.952	0.154	0.002
Female (n^1=10; n^2=8)	0.890	0.013	0.026

dief /d i f/ 'thief'	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.778	0.038	0.004
Female (n^1=10; n^2=8)	0.866	0.805	< 0.001

ziek /z i k/ 'ill'	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.737	0.546	0.101
Female (n^1=10; n^2=8)	0.118	0.507	< 0.001

schip /s x ɪ p/ 'ship'	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.044	0.039	0.598
Female (n^1=10; n^2=8)	0.001	< 0.001	0.784

sik /s ɪ k/ 'goatee'	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.105	0.192	0.146
Female (n^1=10; n^2=8)	0.020	0.005	0.553

pit /p ɪ t/'kernel'	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.020	0.181	0.320
Female (n^1=10; n^2=8)	0.004	< 0.001	0.593

kist /k ɪ s t/ 'chest'	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.008	0.157	0.241
Female (n^1=10; n^2=8)	0.007	< 0.001	0.197

muur /m y r/ ‘wall’	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.890	0.579	0.588
Female (n^1=10; n^2=8)	0.053	0.735	0.744

stuur /s t y r/ ‘steering wheel’	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.079	0.810	0.973
Female (n^1=10; n^2=8)	0.582	0.787	0.803

vuur /v y r/‘fire’	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.567	0.708	0.899
Female (n^1=10; n^2=8)	0.487	0.629	0.222

duur /d y r/‘expensive’	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.505	0.558	0.137
Female (n^1=10; n^2=8)	0.703	0.756	0.512

put /p ʏ t/ ‘pit’	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.115	0.002	0.477
Female (n^1=10; n^2=8)	0.021	0.023	0.303

schub /s x ʏ p/ ‘scale’	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.127	0.709	0.796
Female (n^1=10; n^2=8)	0.027	0.001	0.869

kus /k ʏ s/ ‘kiss’	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.105	0.281	0.238
Female (n^1=10; n^2=8)	0.058	0.049	0.538

brug /b r ʏ x/‘bridge’	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.369	0.388	0.201
Female (n^1=10; n^2=8)	0.104	0.062	0.670

2 English production task

key /k iː/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.240	0.654	0.702
Female (n^1=10; n^2=8)	0.006	0.602	0.106

sheep /ʃ iː p/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.318	0.521	0.004
Female (n^1=10; n^2=8)	0.178	0.359	0.625

peace /p iː s/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.380	0.589	0.527
Female (n^1=10; n^2=8)	0.092	0.756	0.584

leaf /l iː f/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.855	0.774	0.318
Female (n^1=10; n^2=8)	0.481	0.916	0.669

spit /s p ɪ t/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.600	0.459	0.806
Female (n^1=10; n^2=8)	0.167	0.003	0.131

fist /f ɪ s t/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.113	0.527	0.236
Female (n^1=10; n^2=8)	0.076	< 0.001	0.274

six /s ɪ k s/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.636	0.931	0.892
Female (n^1=10; n^2=8)	0.993	0.407	0.329

pig /p ɪ g/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.269	0.484	0.129
Female (n^1=10; n^2=8)	0.046	0.072	0.071

book /b ʊ k/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.523	0.993	0.962
Female (n^1=10; n^2=8)	0.584	0.121	0.389

foot /f ʊ t/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.249	0.719	0.570
Female (n^1=10; n^2=8)	0.267	0.208	0.504

hook /h ʊ k/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.966	0.958	0.645
Female (n^1=10; n^2=8)	0.328	0.578	0.999

cook /k ʊ k/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.520	0.789	0.230
Female (n^1=10; n^2=8)	0.008	0.080	0.738

hug	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.009	0.389	0.518
Female (n^1=10; n^2=8)	0.232	0.088	0.441

duck /d ʌ k/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.124	0.136	0.700
Female (n^1=10; n^2=8)	0.737	0.705	0.389

cup /k ʌ p/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.198	0.548	0.198
Female (n^1=10; n^2=8)	0.239	0.381	0.381

bus /b ʌ s/	F1 (p)	F2 (p)	Duration (p)
Male (n^1=2; n^2=4)	0.891	0.459	0.711
Female (n^1=10; n^2=8)	0.019	0.087	0.250

Appendix K. Production results from Verhoeven & Van Bael (2002b)

1 Average F1 and F2 values for Antwerp and East-Flemish productions of Standard Dutch /i ɪ y ʏ/

	Vowels →	/i/		/ɪ/		/y/		/ʏ/	
		F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
Participant group	Antwerp males	290	2245	329	2122	297	1833	347	1747
	Antwerp females	298	2654	326	2600	302	2051	341	2032
	East-Flemish males	259	2188	392	1926	271	1797	415	1504
	East-Flemish females	286	2465	458	2364	282	2033	475	1767

(adapted from Appendix 1 in Verhoeven & Van Bael, 2002b: 19. For the complete reference of Verhoeven & Van Bael, 2002b we refer to section 8)

2 Average vowel duration values for Antwerp and East-Flemish productions of Standard Dutch /i ɪ y ʏ/

	Vowels →	/i/	/ɪ/	/y/	/ʏ/
		Duration (ms)	Duration (ms)	Duration (ms)	Duration (ms)
Participant group	Antwerp speakers	261	128	290	148
	East-Flemish speakers	97	78	193	88

(adapted from Appendix 2 in Verhoeven & Van Bael, 2002b: 20. For the complete reference of Verhoeven & Van Bael, 2002b we refer to section 8)

Appendix L. Production results from Hawkins & Midgley (2005)

Individual and average F1 and F2 values (speaker group aged 20-25)

	Vowels →	/iː/ (heed)		/ɪ/ (hid)		/ʊ/ (hood)		/ʌ/ (hud)	
		F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
Speaker	Speaker 1	261	2282	454	2375	454	1550	592	1242
		288	2304	452	2032	420	1741	668	1163
		307	2373	424	2003	409	1657	690	1202
		328	2416	441	2098	426	1610	654	1105
	Speaker 2	290	2215	368	2105	400	1200	539	1301
		257	2263	345	2089	452	1332	471	1239
		290	2255	338	2161	423	1198	537	1390
		251	2223	390	2079	381	1158	587	1347
	Speaker 3	235	2552	378	2438	421	1222	922	1224
		270	2479	411	2369	424	1150	805	1191
		238	2434	421	2360	446	1252	805	1107
		295	2459	366	2442	436	1174	738	1124
	Speaker 4	295	1962	346	2000	353	1087	500	1224
		291	1996	342	1974	370	1100	620	1107
		309	1983	374	1930	369	1153	587	1207
		269	2045	335	1996	357	1281	572	1218
	Speaker 5	244	2632	414	2311	453	1275	670	1177
		268	2640	454	2256	393	1223	679	1155
		284	2627	389	2278	417	1125	784	1178
		253	2612	415	2191	453	1245	744	1170
Average		**276**	**2338**	**393**	**2174**	**413**	**1285**	**658**	**1208**

(adapted from Table 1 and Table 3 in the Appendix of Hawkins & Midgley, 2005: 195-199. For the complete reference of Hawkins & Midgley, 2005 we refer to section 8)