Non-native vowel perception

The interplay of categories and features

ADAM MICKIEWICZ UNIVERSITY IN POZNAŃ

SERIA FILOLOGIA ANGIELSKA NR 58

Anna Balas

Non-native vowel perception
The interplay of categories and features

POZNAŃ 2018

ABSTRACT. Balas Anna, *Non-native vowel perception. The interplay of categories and features* [Percepcja samogłosek w języku nienatywnym. Oddziaływanie kategorii i cech] Adam Mickiewicz University Press. Poznań 2018. Pp. 184, Seria Filologia Angielska nr 58. ISBN 978-83-232-3351-0. ISSN 0554-8144. Text in English with a summary in Polish.

The book *Non-native vowel perception: The interplay of categories and features* is devoted to vowel perception in the second, third and foreign language by Polish advanced learners English, French or Dutch as the second and third language in a formal classroom instruction setting. So far it has been assumed that non-native sound perception is based on assimilation to the first language categories or new category formation. The present book hypothesizes that also individual phonetic features, which the learner is familiar with, and the lack of reaction to unknown features, play a role in speech perception. Numerous perceptual tests have been used to examine the development of English vowel perception and the perception of Dutch and Turkish vowels by the learners of English, French and Dutch. The aim of the first study was to test which features ease perception development. The aim of the second study was to examine whether and, if so, to what extent, the familiar phonetic features influence non-native perception. The studies have confirmed the main impact of categories, which act as magnets, but they have also shown the role and the hierarchy of phonetic features in non-native vowel perception.

KEY WORDS: vowel perception, L2, L3 and foreign language, second language, third language, phonetic category, phonetic feature

Anna Balas, Faculty of English, Adam Mickiewicz University, al. Niepodległości 4, 61-874 Poznań, Poland; email:abalas@wa.amu.edu.pl

Reviewer/Recenzent
dr hab. Andrzej Porzuczek

Publikacja powstała w wyniku realizacji projektu badawczego o nr UMO-2015/17/B/HS2/01246 pt. „Przekształcenia percepcji mowy: oddziaływanie kategorii i cech” finansowanego ze środków Narodowego Centrum Nauki. This publication has been a result of a research project nr UMO-2015/17/B/HS2/01246 “Perceptual reorganization of speech: The interplay of categories and features” financed by the National Science Centre, Poland.

Cover design/Projekt okładki: Marcin Gruchociak
Przygotowanie okładki do druku: Pracownia Wydawnicza WA UAM
Typsetting and formatting/Skład i formatowanie: Pracownia Wydawnicza WA UAM

ISBN 978-83-232-3351-0
ISSN 0554-8144

WYDAWNICTWO NAUKOWE UNIWERSYTETU IM. ADAMA MICKIEWICZA W POZNANIU
61-701 POZNAŃ, UL. FREDRY 10
www.press.amu.edu.pl
Sekretariat: tel. 61 829 46 46, fax 61 829 46 47, e-mail: wydnauk@amu.edu.pl
Dział Promocji i Sprzedaży: tel. 61 829 46 40, e-mail: press@amu.edu.pl
Wydanie I. Ark. wyd. 10,8. Ark. druk. 11,5
DRUK I OPRAWA: VOLUMINA.PL DANIEL KRZANOWSKI, SZCZECIN, UL. KS. WITOLDA 7-9

Contents

Acknowledgements

I would like to thank Katarzyna Dziubalska-Kołaczyk, Magdalena Wrembel, Kamil Kaźmierski, Jarosław Weckwerth, Grzegorz Michalski, Geoff Schwartz, Jolanta Szypra-Kozłowska, Andrzej Porzuczek, Arkadiusz Rojczyk, Dafydd Gibbon, Michael Tyler, Ocke-Schwen Bohn, Ulrike Gut, Cécile Fougeron and members of the audience for either their advice on the design of the study or the manuscript or comments after my talks at the Faculty of English in Poznań or at the International Symposium on Monolingual and Bilingual Speech 2015 and 2017, Poznań Linguistic Meeting in 2015, the Laboratoire de Phonetique et Phonologie at the Université Sorbonne Nouvelle Paris 3 in 2015, New Sounds 2016 in Aarhus, Accents 2016 in Łódź, the LSA 2017 Annual Meeting in Austin, Texas, L3 Workshop in Poznań, Approaches to Phonology and Phonetics in Lublin in 2017 and the 50th Annual Meeting of the Societas Linguistica Europaea in Zurich. I am grateful to Hanna Wdowicka for helping me do the statistical analyses. I need to thank Jarosław Weckwerth, Magdalena Wrembel, Anna Matera Klinger, Paulina Zydorowicz, Danuta Wolfram-Romanowska, Dawid Pietrala, Halszka Bąk, Bartosz Brzoza, Nancy Saeys, and Zbigniew Nadstoga for their support in encouraging student participation in the experiments. I am also grateful to the native speakers who recorded the stimuli and participants in the numerous series of the experiments. Finally, I would like to thank my family for their support and encouragement. The project was financed by grant no. UMO-2015/17/B/HS2/01246 "Perceptual reorganization of speech: The interplay of categories and features" from the National Science Centre, Poland.

List of tables

List of figures

Chapter One

Introduction

"To strive, to seek, to find and not to yield"
Ulysses, Alfred Tennyson

Listening is a domain in which we can observe that the human mind is a seeker of patterns. In the case of native speech perception we can appreciate how efficient the seeking of patterns may be and how effortlessly we seem to be processing the speech signal. Speech perception in a non-native language is more challenging: in addition to the normal filtering of environmental effects, the sound categories of the first language (henceforth L1) filter the sounds of the second language (henceforth L2) (Trubetzkoy 1939/69): the L2 sounds, even if there are more of them or they are different, are perceived in terms of L1 categories, or the weight of a given sound feature in L2 is over- or underestimated. Thus far, L2 speech perception has been analyzed in terms of (a) assimilations of those L1 sounds which are similar to L2 sounds or (b) new category formation for markedly different sounds (Flege 1995, Best 1995, Best and Tyler 2007). Pajak and Levy (2014) postulate an important role for selective attention to features in L2 speech perception, which in addition to categories, might be employed in L2 speech perception even if the context for using these features is different.

The present book aims to contribute to the discussion of the *perceptual foreign accent* (Strange 1995: 22) by examining the hypothesis regarding the interplay of categories and features in non-native vowel perception. The focus will be on learners in a formal classroom setting. There are two aspects which will be studied. The first one is a longitudinal study of vowel perception in second language acquisition. Specifically, the development of perception of British English monophthongs will be studied among native Polish listeners, first-year university majors undergoing intensive language and pronunciation training. The second aspect of non-native vowel perception presented here will be the perception of third and foreign language Dutch and Turkish vowels. We will try to capture non-native speech perception in a continuum from developing second-language perception through third language, and finally to foreign language speech perception.

1.1. Theories of speech perception in non-native language acquisition

Several decades of research into speech perception of non-native sounds, its relation to native perception and perception by native speakers of the target language have shown that explaining why certain sounds are easy and others more challenging to perceive is a complex endeavor. This section discusses the L1 phonological filter, markedness, the Perceptual Assimilation Model (Best 1993, 1994a, 1994b, 1995, Best and Tyler 2007), the Speech Learning Model (Flege 1995) with a special emphasis on the feature hypothesis (McAllister, Flege and Piske 2002), the Native Language Magnet (Kuhl 1994, 2000a, Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola and Nelson 2007) and Second Language Linguistic Perception (Escudero and Boersma 2004, Escudero 2005 and 2009).

1.1.1. L1 phonological filter

L2 learners do not start to learn and perceive L2 in a vacuum. Theories of second language speech perception assume that the perception of L2 phonetic segments is heavily influenced by the L1 phonological system. Trubetzkoy (1939/69) claimed that L1 phonology acts as a "sieve" for acoustic differences in the L2 that are not present in the L1. Non-native speech sounds are incorrectly interpreted, because they pass through the filter of native phonology. Weinreich (1963), Brière (1966) and Wode (1977) assumed that L2 learners identify L2 sounds as L1 sounds even if they are different. The Speech Learning Model (Flege 1995) changed the approach to the role of the L1 by introducing the concept of *equivalence classification*.

Michaels (1974) put forward a hypothesis that the relative importance of distinctive features may vary depending on the language, and influence L2 speech perception. Sebastián-Gallés and Soto-Faraco (1999) claimed that children learn to weigh acoustic features of speech optimally for their L1. L2 speech will be perceived in terms of L1-specific feature weightings.

If the L1 acts as a phonological filter then adult listeners find it challenging to discriminate non-native contrasts which are not present in the L1. For example, Japanese listeners distinguish between English /r/ and /l/

at the chance level, as Japanese does not involve this contrast at the phonemic level (Goto 1971, Miyawaki, Strange, Verbrugge, Liberman, Jenkins and Fujimara 1975, Best and Strange 1992).

The concept of phonological deafness was extended beyond the level of segments to new types of suprasegmental and phonotactic deafness (Dupoux and Pepperkamp 2002).

1.1.2. Markedness

Apart from the fact that certain categories might be absent in the L1 and present in the L2, or that features may not play a role in the L1 and be active in the L2, or that features may play a role in different domains in both languages, cross-linguistic transfer might also depend on the markedness/naturalness of a given phenomenon. The Contrastive Analysis Hypothesis (Lado 1957) simply assumed that areas of difficulty might be predicted on the basis of comparison of the L1 with the L2 – what is different will be difficult. The Markedness Differential Hypothesis (Eckman 1977) specified that only those areas which are different and more marked in the L2 will be difficult to acquire. Markedness was defined in the following way: "A phenomenon A in some language is more marked relative to some other phenomenon B, if, cross-linguistically, the presence of A in a language implies the presence of B, but the presence of B does not necessarily imply the presence of A" (Eckman 1981: 211). Dziubalska-Kołaczyk (1990) employed relative markedness in a Natural Phonological model of second language acquisition and Schmid (1997) elaborated on the idea and proposed the Naturalness Differential Hypothesis.

1.1.3. Perceptual Assimilation Model

The Perceptual Assimilation Model (PAM) (Best 1993, 1994a, 1994b, 1995, Best and Tyler 2007) assumes that speech perception depends on recognizing phonological distinctiveness and phonological constancy within a language. When learning a non-native language, learners should shift their attention to higher-order phonetic invariants that distinguish one category from another in that language, and ignore information irrelevant for a given non-native contrast, even it signals a contrast in the L1. The major tenet of PAM is that discrimination varies depending on how contrasting non-native phones are categorized and goodness-rated in terms of native language phonological categories. In the case of Two-Category assimilation, namely when two non-native phones are assimilated to two different native phonological categories, discrimination is expected to be excellent. Discrimination is predicted to be poor, however, in Single-Category assimilation; the L2 phones are heard as equally good or poor versions of the same native phonological category. Alternatively, when two L2 phones are assimilated to the same L1 phonological category, but one of them is judged to be a better exemplar of it than the other, which is called category goodness assimilation, discrimination rates are expected to be worse than in the case of a Two Category assimilation type, but better than in the case of a Single Category. An L2 phone can also be uncategorized in terms of L1 phonemes. This happens when an L2 phone is assimilated to a comparable extent to two or more L1 categories. Such a scenario yields either Uncategorized-Uncategorized contrasts or Uncategorized-Categorized contrasts. Discrimination rates in Uncategorized-Uncategorized contrasts were first hypothesized to vary from poor to excellent, depending on the phonetic similarities between the two L2 phones and to the phonological categories in the L1, but they have now been elaborated on by Faris, Best and Tyler (2016; see below). Discrimination of Uncategorized-Categorized contrasts should be very good, as they cross a category boundary in the L1. Similarly, for non-assimilable phones, predictions are based on their non-speech auditory similarity, though non-assimilable vowels are unlikely to exist as Tyler et al. (2014 p. 6) noticed.

In contrast to previous research on language-specific tuning in speech perception which primarily examined consonants, Tyler et al. (2014) tested American English speakers' perception of six non-native vowel contrasts. Their aim was to see whether non-native vowel perception is governed by the same principles as postulated by PAM and the Natural Referent Vowel framework (Polka and Bohn 2003, 2011). They showed that vowel discrimination depended on assimilation patterns as predicted by PAM. Asymmetries hypothesized by the Natural Referent Vowel framework, however, were found only in the case of Single Category assimilations. This finding was taken to suggest that assimilation types might influence the ways in which peripheral vowels affect vowel perception. Non-native vowel contrasts which cross a phonological boundary, that is Two Category and Categorized-Uncategorized assimilation types, mitigated the effects of vowel peripherality on perceptual asymmetries. Tyler et al. (2014) therefore concluded that peripheral vowels may influence adult non-native vowel discrimination when native phonological distinctions do not interfere, as in the case of Category Goodness, Single Category and Uncategorized-Uncategorized assimilation types.

Second language category formation is assumed to depend on the degree of perceived differences between L1 and L2 phones (Best 1995, Flege 1995, Best and Tyler 2007). In the case of uncategorized non-native phones, the degree of perceptual overlap between the sets of native categories used to assimilate each phone from the L2 pair should then be crucial for predicting category formation. Faris, Best and Tyler (2016) examined Egyptian Arabic listeners' perception of Australian English vowels. Because of vowel repertoire differences between Egyptian Arabic (10 vowels) and Australian English (19 vowels) they found numerous examples of phones which were not assimilated to a single L1 category above a predefined threshold of e.g. 50% or 70%, i.e. uncategorized phones. They showed that the uncategorized phones vary in the way they map onto a native phonological system. When a non-native phone is perceived as similar to one native category, but categorized below a categorization threshold, it is called *uncategorized focalized*. When a non-native phone is perceived as similar to a small set of native categories, the responses are termed *clustered*. In the case of *dispersed responses*, numerous native categories are randomly chosen as targets, suggesting that a non-native phone is not similar to any of the native categories.

In PAM (Best 1995), discrimination accuracy for Uncategorized-Uncategorized assimilation types was predicted to vary from poor to moderate, depending on the similarity of the two non-native phones to the same set of native phones and the similarity of the two phones. Distinguishing between the three uncategorized assimilation types allowed Faris et al. (2016) to draw more precise discrimination predictions for Uncategorized-Uncategorized assimilation types. Under the assumption that each of the uncategorized phones in a pair is assimilated to a different native category or a set of categories, phones in the focalized-focalized assimilation type should be relatively easy to discriminate, followed by focalized-clustered, clustered-clustered, focalized-dispersed, clustered-dispersed, and finally dispersed-dispersed which are predicted to be the most difficult to discriminate.

As the degree of overlap between the categorization of L2 phones to L1 categories also influences discrimination performance (Tyler et al. 2014), Faris et al. (2016) postulate separate assumptions for uncategorized phones which are perceived as similar to the same set of categories in L1. Focalized and clustered contrasts that are assimilated to the same set of L1 categories (i.e. completely overlapping) will be less discriminable than focalized or clustered contrasts assimilated to a different set of native categories (i.e. non-overlapping). Non-overlapping phones will be more easily discriminated because of the perceived phonetic similarity of each non-native phone to a different set of native categories. Focalized or clustered contrasts which partially overlap will be less discriminable than non-overlapping contrasts and more discriminable than the completely overlapping contrasts.

The results of Faris et al. (2016) also inform PAM-L2. As in the case of a *new* L2 phone in the Speech Learning Model (Flege 1995), in dispersed assimilations L1 categories do not interfere, because listeners do not detect L1 category invariants in the L2 phones. This makes it possible to form a new phonological category. In the case of focalized and clustered assimilations, the degree of overlap with contrasting L2 phones will affect the chances of forming a new category. A new L2 category will likely be formed if the focalized or clustered phone does not overlap with any other non-native category.

The major issues to now be solved within PAM are the following: to evaluate discrimination predictions for uncategorized phones and the influence of perceptual overlap on discrimination rates as well as *new* category formation and to replace the use of an arbitrary categorization threshold (Faris et al. 2016). The present study may not contribute to these aims, but it uses the three uncategorized assimilation types distinguished by Faris et al. (2016) to trace the perception of Dutch and Turkish vowels by the three groups of Polish learners of Dutch, French and English.

1.1.4. Speech Learning Model

In the Speech Learning Model (Flege 1995), L2 sounds are classified in relation to L1 sounds in the following terms: *identical*, *similar* and *new* (Flege 1987a, b, 1997, Bohn and Flege 1997, Flege, Bohn and Jang 1997). An identical L2 sound is produced authentically due to positive transfer. A similar L2 sound has a counterpart in the L1, though it differs in some way. A new L2 sound is not acoustically or perceptually similar to any of the L1 sounds and therefore learners should easily notice that it is not a realization of a native category. Flege (1987b, p. 48) gives an example of the French vowel /y/, which has no counterpart in English and should be recognized a new category by L1 English learners of French. Bohn and Flege (1992) assume that an L2 vowel is a new vowel for L2 learners if most of its realizations occur in the acoustic phonetic vowel space which is empty in the L1, while for learners with a large L1 vowel inventory, few of the L2 vowels could be new. Rochet (1995, p. 390) disagreed with these claims and suggested that new and similar L2 sounds are those L2 phones which are perceived to be so by L2 learners and therefore the concept of uncommitted space should only be considered in perceptual, rather than in acoustic terms. He shows that hardly any L2 phones can be considered new, i.e. not belonging to the same category as any of the existing L1 sounds. Rochet (1995) proposed that categories can extend to the limits of the next categories, so that L2 phones that are not found in the L1 inventory are nevertheless perceived as belonging to a native category or recognized as nonlinguistic sounds (clicks being an example here).

The major contribution of the Speech Learning Model lies in the formulation of equivalence classification in the domain of second language acquisition. Previously, the L1 was seen as a simple filter blocking any L2 sound properties that do not occur in the L1 (cf. section 1.1.1). Equivalence classification is not an auditory or phonological filter for subphonemic differences between the L1 and L2 sounds (Flege 1997, p. 16). Equivalence classification is defined as a basic cognitive mechanism, which enables humans to perceive constant categories among physical exemplars characterized by the inherent sensory variability (Flege 1987a). Equivalence classification is vital in L1 acquisition, because it permits children to identify phones spoken by different people and in different phonetic contexts as belonging to the same category. This mechanism, which is so important in L1 acquisition, prevents older children and adults from using auditorily accessible acoustic differences between the L1 and L2 sounds. As a result, when hearing L2 sounds older children or adults seek constancy in the sensory information, are likely to process L2 phones as L1 sounds, and ultimately speak the L2 with an L1 accent. According to the Speech Learning Model, one's phonetic system may be influenced by all audible acoustic differences between the L1 and L2 sounds, including the ones which are not perceptually available (Flege 1987a, p. 16). L2 learners may not perceive the differences between the L1 and L2 sounds, not because they are unable to detect them, but because of equivalence classification.

The existing phonetic categories may be altered to better accommodate the acoustic properties of similar L1 and L2 sounds. Even when equivalence classification has blocked new category formation, learners may eventually perceive differences between the L1 and L2 sounds. The process is generally slow and limited by influences from both languages. This happens in the case of a diaphone or merged category characterized by bidirectional cross-linguistic influence. Such a scenario is given as an explanation of the fact that Italian learners of English pronounce English voiced plosives with compromised VOT values in both Italian and English. The assumption that a phonetic system develops even in adulthood as a result of the L2 input differentiates the Speech Learning Model from previous theories.

Flege's Speech Learning Model (Flege 1995) hypothesizes that the smaller the perceived phonetic distance between an L1 and an L2 vowel, the more challenging it will be for the learner to establish a new category for the L2 sound. The Speech Learning Model does not claim that equivalence classification is an auditory or phonological filter for subphonemic differences between L1 and L2 sounds (Flege 1997). Rather, the Speech Learning Model states that one's phonetic system may be influenced by all audible acoustic differences between L1 and L2 sounds, "even those that are not available perceptually" (Flege 1987a, p. 16). Perception of the differences may be blocked by equivalence classification, not necessarily by the inability to detect auditory differences.

Accurate perceptual targets are important in the Speech Learning Model because they control the sensorimotor learning of the L2 sounds (Flege 1995, p. 238). Long-term memory representations called phonetic categories include specification of the sounds of speech. L1 and L2 categories are compared at the level of position-sensitive allophones, which is not as abstract as the phoneme level. Phonetic categories can evolve as a result of the L2 input when the L2 phones are identified as belonging to a given L1 category. Bilinguals have one phonological space for all the L1 and L2 categories they have formed and they will try to maintain the contrast between them. Learners are more likely to discern the sounds and form a new category when they perceive an L2 sound as phonetically dissimilar to the L1 sounds. Even when a new category has been formed for an L2 sound, its specifications can differ from a monolingual's, if the category is altered to maintain contrast with other categories in a common L1-L2 phonological space or if the learner uses other features or feature weights as compared to the monolingual.

1.1.5. Feature hypothesis and its extension

The fifth hypothesis in Flege's (1995) Speech Learning Model states that category formation for an L2 sound may be blocked by equivalence classification. McAllister, Flege and Piske (2002) further developed that idea and proposed the feature hypothesis: the L2 features which are not used as contrastive ones in the L1 will be difficult to perceive for the L2 learners. Consequently, forming an L2 category may be blocked if the category crucially depends on a feature which does not have a distinctive function in the L1. The results of the study by McAllister et al. (2002) demonstrated that the acquisition of short and long vowels in L2 Swedish depended on the status of the duration feature in the subjects' L1s. Learners with L1 Spanish, which does not use a duration feature at all, found the contrast the most challenging. Their results were only slightly worse than the results for English subjects, who use duration only as a secondary cue in their L1. Estonian learners, whose L1 uses vowel duration contrastively, were significantly more successful.

The feature in the Speech Learning Model (Flege 1995) and feature hypothesis (McAllister et al. 2002) cover both distinctive phonological features and phonetic non-distinctive features.

Aoki and Nishihara (2013) extended the hypothesis to cover the difficulty in acquiring features which are not active in previously acquired languages. In their experiment, they tested VOT values as produced by L1 Japanese learners of L2 English (more than six years of instruction), some of whom had learned Chinese as their L3 for one to four years. Japanese does not have aspirated consonants, while aspiration in English is widespread though it has an allophonic status only, whereas in Chinese it has a distinctive function. Subjects were asked to read carrier sentences in English with words containing aspirated and unaspirated stops. The results revealed that the VOT difference was sufficient in the case of an English control group and Japanese learners of English who had Chinese as their L3, but not in the case of Japanese learners of English who had not learned Chinese. Therefore, Aoki and Nishihara (2013) extended the feature hypothesis to claim that a non-native feature is easy to acquire if it is used in any acquired language, regardless of the bilingual or monolingual environment, the order of acquisition of non-native languages or the level of proficiency in the source and target language. They did not comment further on the distinctive function of a given feature, but it seems that VOT differences functioning as secondary cues in English were not as robust as phonemically distinctive VOT differences in Chinese.

1.1.6. Native Language Magnet

Although initially developed to account for the transition from auditory to language-specific perceptual processing, the Native Language Magnet (Kuhl 1994, 2000a) and the expanded Native Language Magnet model (Kuhl et al. 2008) claim, with regard to L2 speech acquisition, that the most often activated phonetic representations begin to act as magnets for other members of the category in the L1 but at the same time they reduce foreign language phonetic abilities. The scenario of early speech perception development, which accounts for the way infants form native phonetic categories thanks to ambient language experience, is as follows. Initially, due to general auditory processing mechanisms, infants are able to differentiate all sounds of human speech (Kuhl 1991b). The first phase lacks a speech specific mechanism. Secondly, infants develop sensitivity to the distributional patterns and exaggerated cues of infant-directed speech. At the end of phase two, the perception of native language phonetic cues is enhanced, whereas discrimination of non-native contrasts is reduced. Experience warps perception and produces a distortion which causes desensitization near category modes and boosts sensitivity at category boundaries (Kuhl 1991a, Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Ketterman and Siebert 2003). With growing experience, prototypes based on the representations which are most often activated start to act as perceptual magnets for the sounds identified as belonging to a given category and they increase the perceived similarity between realizations of the category. In phase three, the enhanced speech perception skills further boost three skills which infants need for word acquisition: the detection of phonotactic patterns, the detection of word-like units, and the association between sound combinations and objects. Awareness of phonotactic distinctions is improved by learning phonetically similar words. Lastly, experience with ambient language results in relatively stable neural representations, which are not easily shifted by new input. Short periods of listening to a non-native language do not automatically form new neural structures. The perceptual magnet effect, which facilitates native language processing, reduces sensitivity to L2 contrasts, as perceptual mapping in L1 creates "a complex network, or filter, through which language is perceived" (Kuhl 2000a, p. 11854). Perceptual attunement to L1 categories interferes with L2 speech perception because one cannot separate L1 and L2 mappings, or categories, and L1 category mappings influence L2 category processing.

1.1.7. Second Language Linguistic Perception Model

The Second Language Linguistic Perception Model (Escudero and Boersma 2004, Escudero 2005 and 2009) was designed to describe, predict and explain L2 perception from the time individuals are not yet learners, through the time when they begin to learn the L2 and the developmental state, to the final stage in learning the L2. The initial state is claimed to be a copy of the L1 system, which determines the individual's learning tasks in L2 development. According to the Second Language Linguistic Perception Model, the initial state is defined on the basis of a detailed acoustic comparison of the production of native and second language sounds.

The learning task depends on the transformations which need to take place between the initial state and the target L2 perception. The model's predictions resemble those of the Perceptual Assimilation Model (Best 1995, Best and Tyler 2007) – both models are interested in contrasts, rather than in individual categories. An L2 contrast is considered new when the two L2 sounds are categorized to the same native language category. In the case of such a new scenario, the learner's task is to create a new category for one of the L2 sounds or split the existing category. When two L2 sounds are heard as two separate L1 categories, we have a similar scenario, in which learners are not expected to form new categories. Their task then is to shift the boundaries of categories to accommodate the L2 sounds. Category shift is considered to be an easier endeavor than category formation or category split. The third option is a subset scenario, in which an L2 sound is heard as more than one native category. The multiple category assimilation should be less challenging than the new scenario, because no new contrasts should be formed in the L2, and discrimination is expected to be high. This scenario may, however, be challenging when an L1 contrast is active during the acquisition of an L2 without this contrast, which may lead to over-differentiation at the word level.

The Second Language Linguistic Perception Model assumes that perception development can be based on either distributional learning or on lexical- or meaning-driven learning. Escudero and Williams (2014) showed that distributional learning, which is a statistical learning mechanism based on the relative frequency of exposure to auditory stimuli that form a continuum along an acoustic dimension covering two L2 categories, proved beneficial for non-native sound discrimination over 12 months. Bimodal distribution, where the extreme tokens in the continuum are most frequently presented, or enhanced bimodal distribution, where the F1 and F2 were exaggerated, were more effective than unimodal distribution, where the tokens from the middle of the continuum were most frequently presented.

The second type of perception development is meaning-driven (van Leussen and Escudero 2015). Perceptual and lexical learning are intertwined. Lexical processing can either be bottom-up or interactive, i.e. profiting from lexical information. Computational simulations in van Leussen and Escudero (2015) showed that both strategies lead to adequate recognition in the L2, but the simulated learners using an interactive processing strategy needed less exposure to L2 words.

The end state in the Second Language Linguistic Perception Model is the ultimate attainment learners can reach, but the attainment may not or does not need to match the L2. How quickly a learner reaches the final state depends on the specific learning tasks (new, similar or subset scenarios) and on richness of the input. The model predicts that L1 perception may only be influenced by the L2 input, if there is little L1 input. At this stage the learners are hypothesized to have two different grammars and activation modes which allow for optimal native perception and optimal L2 perception.

1.2. Features

Phonologists define a distinctive feature as a feature representing a minimal contrast between two sounds (Daunmu 2016: 8). A contrast is then defined as the phonetic difference between a pair of sounds that can distinguish words in a language. Such an approach has a clear advantage – those phonetic features which are not contrastive, such as, for example, differences in vowels as produced by adults and children or males and females, are not represented by features. Features are often also defined by sound classes, in such a way that a minimal similarity that members of a sound class have in common is a feature and a sound class is a group of sounds which undergo the same phonological processes. As Daunmu (2016: 10) however notices it is not always easy to agree which sounds belong to a sound class and what feature they share.

In phonetics features represent the phonetic properties of sounds. Those properties refer to articulatory gestures, acoustic or perceptual effects. Daunmu (2016: 12) notes that it is sometimes challenging to determine which phonetic properties are relevant or how many degrees of a feature we should distinguish. Phonologists have traditionally used binary features

(Jakobson, Fant and Halle 1951, Chomsky and Halle 1968), but the use of privative, or one-valued features has also been postulated (Lombardi 1991, 1995 and Steriade 1995)[1]. Phoneticians are more likely to assume that phonetic features are gradient. Daunmu (2016) also gives examples to show that phonetic features may need to be reconsidered in the light of information available from contrast and sound classes. He also observes that although according to Halle (1962 and 1995) features derived from the three domains, i.e. contrast, sound classes and phonetic properties should be in agreement with one another, we are still looking forward to prove it is the case.

1.3. Features and categories in non-native language acquisition

The ontological status of features and categories assumed in this project can be approached from many perspectives. There is a diversity of terminology in the relevant literature. Flege (1995) talks about phonetic features and phonetic categories. Best (1995) assumes higher-order phonetic invariants which are necessary and sufficient for distinguishing one category from other categories that contrast with it in the native language phonological space – so the categories are phonological. Pajak and Levy (2014) claim that "in addition to learning individual phonetic categories learners are posited to make higher-order generalizations about the general properties of the set of those categories, which are also called abstract phonological principles". Pajak and Levy (2014) also ask questions about the abstractness of these phonological principles: "do they apply only in cases when the corresponding contrasts are phonetically similar (as is the case with vowels and glides)? Or are they more abstract, applying regardless of the degree of phonetic similarity between known and non-native contrasts?" (p. 155), favoring the second option. The present author's view, along Natural Phonological lines (for more recent publications see Stampe and Donegan 2009 and Donegan 2015), is that acoustic-phonetic features can acquire phonological status when they are used to distinguish contrasts, i.e. phonological categories in a language. According to this view, a phonetic feature might become phonological in the phonological system of a given language.

---

[1] See Backley (2011) for a discussion of the rationale for binary or privative features.

Importantly, Natural Phonology accounts for gradience in the speech signal by assuming that multiple factors interact in speech production. The interacting forces encompass feature combinations, sound sequences, position within a word, prosody and speech rate (Stampe and Donegan 2009, Donegan 2001). No additional gradient specifications are required, because phonetic gradience stems from simultaneous and sequential feature combinations, prosodic patterns and process application.

The challenge for second language speech perception is seen in the redirecting of attention to features which may be unimportant or non-existent in the L1 system, but which signal phonological distinctions in the L2. Learners should also disregard phonetic differences which may be crucial in the L1 but do not signal a contrast in the L2. In the initial stages of L2 acquisition, L2 phonological categories are assimilated to L1 phonological categories, where by a phonological category we understand a bundle of acoustic/phonetic features defining a category. More refined L2 speech perception entails taking out a feature listeners are familiar with from the L1 and using it in a different context in the L2 (cf. Pajak and Levy 2014) to construct a phonological category which is appropriate or at least more appropriate for the L2. The proposed project focuses on acoustic-phonetic features, which, when used to distinguish between vowel categories, become phonological.

The big questions which accompany second language speech research concern the nature of the new category together with all sub-questions. How can a category be altered due to non-native language acquisition? Under which circumstances do learners really form new categories (as in the case of identification of /r/ and /l/ by Japanese listeners in the study by Flege, Takagi and Mann (1996))? Do new categories exist in the common phonological space, as the Speech Learning Model (Flege 1995) claims? Do we create separate phonological spaces for each of our languages, as the Second Language Linguistic Perception Model (Escudero 2005) claims? Does a bilingual listener have the same potential as two native speakers of the respective two languages? It is useful to distinguish between a prototypical native speaker of one language in a generally monolingual society, a non-prototypical native speaker, i.e. a bi- or multilingual who uses more than one language on regular basis (Dziubalska-Kołaczyk 2016) and a non-native speaker who is a learner (who may eventually become a non-prototypical native speaker) (Dziubalska-Kołaczyk 2016)? What are the ways of managing categories which come from two or three separate language systems? There are no straightforward answers, not even research designs which could allow researchers to ultimately resolve the complexities, but many detailed phonetic studies can bring us closer to understanding the perception of non-native speech.

1.4. Vocalic systems

One could simply list L1 and L2 vowel categories. We can, however, also trace the common properties of vowels in the L1 and L2 and try to identify what distinguishes vowels in the L1 from vowels in the L2. Information gained on the basis of L2 perception studies can shed light on the interactions between categories, the features that the two sounds in question have in common, and the features that distinguish the categories. One of the interesting issues will be that some features are robust and easy to perceive, while others fail to exert any influence on perception.

Vowels are conventionally described using three parameters: vowel height, backness and lip-rounding (Maddieson 1984:123). The UCLA Phonological Segment Inventory Database (henceforth UPSID), which is a representative sample of 317 languages, recognizes five different heights: high, higher mid, mid, lower mid and low. On the front/back dimension vowels are classified as front, central or back. Finer differences in the case of both dimensions can be resolved by representing the vowel as non-peripheral. Crothers (1978) states that the number of height distinctions is usually equal to or greater than the number of backness distinctions. Maddieson (1984: 124) makes the following observations regarding the distribution of vowels in the vowel space. Mid vowels are slightly more frequent than high vowels (40.5% vs. 39% respectively). Low vowels are the least common (20.5%). Front vowels are slightly more numerous than back vowels (40% vs. 37.8% respectively), whereas central vowels amount to only 22.2%. Unrounded vowels are substantially more common than rounded vowels (61.5% vs. 38.5%). Maddieson (1984) further analyzes the interactions of vowel height, backness and lip rounding. Front vowels are predominantly unrounded (94%), while back vowels are mostly rounded (93.5%). Low vowels tend to be central (75.1%) and central vowels tend to be low (69.4%). High front vowels are more common than high back vowels. Provided the position of the lips is unmarked (i.e. unrounded in the case of front vowels and rounded in the case of back vowels), among mid vowels back vowels are more frequent than front vowels. If the position of the lips is marked, the reverse is true: front rounded vowels are more frequent than back unrounded vowels. There is only one case of a rounded nonback low vowel in the UPSID.

Maddieson (1984) distinguishes between counting the number of vowels per language and counting the distinctive vowel qualities, based on the basic parameters of height, backness and rounding, disregarding differences in duration, nasality or laryngealization. The latter may be more adequate for indicating how intensively the basic vowel features (height, backness and rounding) are exploited (cf. Clements 2003). Over 60% of world languages, including Polish, have between five and seven vowel qualities. Languages that have more than 10 vowel qualities are very rare.

The more vowel quality contrasts a given language has, the more likely it is that it will have vowel length contrasts. The employment of both durational and quality differences encourages the distinctiveness of vowel contrasts (for example, the lengthening of /æ/ in English should perhaps distinguish it from /e/).

Chapter Two

Selective attention to features

2.1. What in L2 speech perception cannot be explained by filtering through an L1 phonetic category inventory

Motivation for Pajak and Levy's (2014) study lay in the observation that certain phenomena in L2 speech perception could not be explained with reference to native contrasts and their phonetic characteristics. Namely, Bohn and Best (2012) found that Danish, French and German listeners exhibited better-than-native discrimination of English /w - j/ contrasts. These results were found intriguing because nonnative listeners performed better at discriminating English contrasts than English listeners did, and, perhaps more surprisingly, Danish and German listeners do not even have /w/ in their native inventory. The authors ruled out the possibility that increasing amounts of spoken English experience could have led to the unpredictable near-ceiling and continuous discrimination of English /w - j/ contrasts, as in such a case the discrimination rates of non-native listeners could merely approach the levels of discrimination of native English listeners and would not exceed them. Previously, Hallé, Best and Levitt (1999) linked better discrimination rates for English approximants by French rather than American and Japanese listeners to the fact that French has a bigger inventory of approximants (/r, l, w, j, ɥ/) than English (/r, l, w, j/) or Japanese (/r, w, j/) and therefore French listeners may be more sensitive to approximant contrasts. The results of the study by Bohn and Best (2012) refuted that hypothesis. Danish and German listeners, whose native approximant system has only three categories in a syllable-initial position (/r, l, j/), while Danish has more allophonic approximants in medial position, discriminated American English /w - j/ comparably to French listeners and decidedly better than native American English listeners. This meant that structural phonological and allophonic differences related to approximants between the three languages could not account for the results. Bohn and Best (2012) noticed, however, that what French, Danish and German share, as opposed to English and Japanese, is a systematic distinction between front rounded and unrounded vowels. The distinctive, contrastive lip rounding feature for vowels is used in French, Danish and German, but is absent both phonologically and phonetically from English and Japanese vocalic systems. Considering the fact that /j/ and /w/ are semivowels, namely short, non-syllabic versions of the unrounded /i/ and rounded /u/, Bohn and Best (2012) proposed that high sensitivity to lip rounding distinctions in native vowels allows listeners to discriminate a corresponding non-native approximant contrast, if that contrast is distinctively differentiated by lip rounding. They also suggested that both PAM and SLM should be amended to account for the effects on cross-language approximant perception of the listeners' native vowel systems, i.e. not only should these models encompass the influence of native categories and their phonetic characteristics on cross-language perception, but they should also account for the role of phonological principles in the native language system.

2.2. Selective attention to features in non-native speech perception

Non-native speech perception can be interpreted as a process involving "learned patterns of selection and integration of those acoustic properties of speech stimuli that are phonologically relevant in the native language" (Strange and Shafer 2007). Recently, Pajak and Levy (2014) postulated the important role of selective attention to features in L2 speech perception. Selective attention to features may account for problems in discriminating the contrasts that differ along unattended dimensions, because L1 perceptual weighing of the relevant acoustic and articulatory cues will influence L2 perception. On the other hand, non-native speech perception might be fostered if the cues that are relevant for the L1 can be used for discrimination in the L2, even if the enhancement should apply in different acoustic and phonetic contexts in the L2. The idea that cues are reused dates back to Clements (2003), who argued that languages reuse a limited set of phonetic features for signaling multiple contrasts. Consequently, it seems that it is more economical for the learner to enhance sensitivity to features which can be reused in numerous contexts rather than enhance sensitivity to specific L1 categories. Pajak and Levy (2014) conclude that speech perception is mediated not only by L1 phonetic categories, but also by more general phonological principles. The issue now is to determine what exactly these phonological principles are and how they influence speech perception.

Pajak and Levy (2014) tested the discrimination of non-native consonant duration contrasts by naïve listeners with different L1s. They considered length to be a suitable testing ground for selective attention to features as both vowels and consonants can employ length-based contrasts. The subjects were recruited from the following language backgrounds: Korean, which uses length contrastively for both vowels and consonants; Vietnamese, where there is contrastive vowel length; Cantonese, in which vowel length is a secondary cue that changes alongside vowel quality; and Mandarin Chinese, where length does not play a role. They treated the informativity of a cue as a continuum, where the cues employed for distinguishing between phonological contrasts would be the most informative ones, while the secondary and allophonic cues would be less informative, but still able to influence listeners' perceptual sensitivity. The following hypotheses were formulated regarding the performance in an AX discrimination of short and long consonants task. If there is general enhanced sensitivity to any cues which are informative in the L1 in any context, then Korean, Vietnamese and Cantonese listeners should perform better than Mandarin speakers. If the degree of cue informativity is vital, then the results should be gradient, depending on the length cue status in the L1: Korean subjects should outperform both Vietnamese and Cantonese subjects, and Mandarin subjects should have the lowest scores. If, in turn, selective attention to features applies only to acoustically or articulatorily similar segments, then only Korean listeners should show good discrimination, as only Korean uses contrastive length for consonants, whereas other groups should have equally poor results, as none of the other languages employs consonant length. In addition to consonant stimuli varying in length, the experiment also included control items with Polish sibilants in place of articulation contrasts (alveolo-palatal and retroflex fricatives and affricates) which are similar to alveolo-palatal and retroflex consonants existing in Mandarin. Consequently, Mandarin listeners were expected to outperform other groups on sibilant discrimination. The stimuli in the study consisted of Polish nonce words with consonant length contrasts and alveolo-palatal contrasts.

The results for length contrasts were very robust. Korean, Vietnamese and Cantonese listeners outperformed Mandarin listeners. There was no significant difference between Korean and Vietnamese subjects, but there were significant differences between Korean and Cantonese subjects and between Vietnamese and Cantonese subjects. To summarize, Korean and

Vietnamese subjects outperformed Cantonese subjects, who in turn outperformed Mandarin subjects. The observed pattern did not exactly confirm any of the hypotheses, which predicted that either all subjects with length contrasts in their L1s would perform equally better than Mandarin subjects or gradiently better than Mandarin subjects, depending on whether their L1 uses length for vowels only (Vietnamese and Cantonese) or for both vowels and consonants (Korean). The results, however, showed that sensitivity to the length cue was related to the informativity of the cue in the L1. Perceptual sensitivity to length is boosted most when the L1 uses it as a contrastive feature, as is the case in Korean or Vietnamese. When length is only a secondary cue to a phonemic contrast, as in the case of Cantonese, perceptual sensitivity is raised to a lesser extent.

This author notices two issues which require further investigation. The first concerns the role of bilingualism in selective attention to features and sensitivity to secondary cues. All the participants in Pajak and Levy's (2014) study were bilingual in English and English uses length as a secondary cue for its tense-lax or free-checked vowel contrasts. Superficially, its phonological role seems to be similar to or only slightly less important than the role of vowel length in Cantonese, where vowel contrasts are based on length and quality differences. There is one Cantonese vowel pair /ɐ-aː/ where the contrast is based primarily on length with small qualitative difference (Zhang 2011) and its vowels are not in complementary distribution. If L1 Cantonese L2 English listeners were more sensitive to length differences than L1 Mandarin L2 English listeners, it would be interesting to know whether this difference was only due to a subtle difference in the role of length in the phonology of Cantonese and English, or whether the L2 is not capable of exerting influence on perception, in a way similar to the one exerted by the L1. Then the question about the role of L2 experience in the perception of non-native contrasts and selective attention to features arises. The second issue is related to the robustness of secondary cues for vowels. As Pajak and Levy (2014) showed, reliance on features which are contrastive in the L1 crucially influences perception, whereas the impact of secondary cues is weaker. Both the role of the L2 and secondary cues will be further investigated here.

2.3. Selective attention in other domains

Pajak and Levy (2014) reported that listeners can use a phonological principle from one class of phonemes in the L1 to perceive another class of phonemes in the L2, even in a different context. Listeners with an L1 which distinguishes between short and long vowels were shown to be able to perceive the difference between short and long consonants in the L2, in spite of the fact that their L1 did not use consonant duration contrastively. Ong, Burnham, Stevens and Escudero (2016) demonstrated that distributional learning of lexical tones and musical pitch by naïve learners leads to cross-domain transfer. Thus, acquiring either speech or musical items distinguished on the basis of pitch differences results in increased sensitivity to pitch in the other domain.

2.4. Why should we try incorporating selective attention to features when accounting for non-native speech perception?

The major problem in speech perception research in second language acquisition is that the existing models (Flege, 1995, Best 1995 and Best and Tyler 2007) do not have much predictive power. This book aims at making predictions as to how a given L2 sound might be perceived by a given L1 speaker more precisely by supporting current category-based predictions with the role of features which interact with major assimilatory or magnet forces of categories.

So far, L2 speech perception research has focused on categories; i.e. which L2 sounds are assimilated to which L1 sounds or when and how a new category can be formed. Recently, Pajak and Levy (2014) have postulated focusing on the role of features in L2 speech perception. Actually, it is surprising that this attention to phonological features has appeared so late, in the light of research on L1 perception, where phonetic detectors have been claimed to operate on features across categories, and the claims of Major's Ontogeny Phylogeny Model (Major 2001) about the role of processes in second language acquisition and Natural Phonological processes which affect a given feature (Stampe 1969, Donegan and Stampe 2009; see also Donegan 2001 and Balas 2009 for accounts of perception). Nevertheless, the role of features has so far been neglected in phonetic studies on L2 perception. The idea of selective attention to features in L2 is worth examining, because the existing and predominant theories do not fully account for L2 speech perception – we can interpret the results using SLM (Flege 1995; Best 1995), but we cannot predict how a given sound will be perceived. By taking into account selective attention to features (Pajak and

Levy 2014) or feature combinations, as this project suggests, we could arrive at more precise predictions concerning L2 perception.

So far, research has concentrated on the comparison of individual sound categories in L1 and L2. Pajak and Levy (2014) postulate examining features as entities whose familiarity contributes to the successful perception of L2 phonemes. This author hypothesizes that these may not necessarily be individual features which ease L2 perception, rather they are configurations of features. For example, Polish learners have many difficulties perceiving English high back vowels (Bogacka 2004), although on the basis of the L1, they are familiar with various degrees of lip rounding and higher and lower back vowels. The challenge is that Polish does not have mid-high back-centralized vowels. So, in fact, it is a combination of features such as partial lip-rounding, centralization and retraction of a back vowel, which is challenging, not the individual categories such as vowel height and vowel centralization.

Also, L2 learners' perception of unfamiliar non-native contrasts has not been thoroughly studied yet. Research on foreign language perception will allow one to distinguish between experience-tuned vs. universal phonetic sensitivities. Studying subjects with a common L1 and different L2s should also allow for tertium comparationis: subjects will have a common repertoire of L1 sounds to refer to in forced-choice categorization and similarity rating tasks, yet their experience with different phonological systems (due to different L2s) should also emerge.

"If we are interested in *phonetic and phonological categorization processes*, we must design studies that begin to reflect real-world stimulus and task constraints in language processing (including those of the language classroom and the L2 work environment) while maintaining experimental control and rigor." (Strange and Shafer 2008, emphasis mine). This book focuses on advanced learners who are formally taught to speak (a) non-native language(s). The existing research concentrates on naturalistic settings in L2 acquisition. The subjects examined within this project might show different patterns of acquisition and their results will reveal how effective formal instruction is.

The present thesis proposes that these phonological principles or processes affecting one feature at a time combine to determine sensitivities to the incoming speech signal in particular ways. Therefore, even if it is assumed after Clements (2003) that a language reuses features, the present project proposes that these phonological principles or processes affecting one feature at a time combine to determine sensitivities to the incoming speech signal. The beneficial effects of two combined features on L2 speech perception should be greater than the use of two separate, unrelated features. This means that reusing a feature such as vowel centralization will be more efficient in the area of the vowel space where the L1 uses this feature, i.e. Polish might centralize the short lax mid-high front centralized vowel and a similar English vowel might be easily perceived. But we cannot assume that the ease of perception will equally well apply if the process of centralization should apply for a high-mid back centralized vowel where L1 Polish does not boast a selection of vowels.

2.5. Research hypothesis

The proposed hypothesis is that speech perception, as reorganized in the course of L2 acquisition, is an interplay of categories and features. It is hypothesized that L2 speech perception is not only governed by similarity or dissimilarity to L1 speech sound categories, as postulated by two prevailing models in the field (Flege 1995, Best 1995, Best and Tyler 2007) but that a vital role is also played by sensitivity to features used in L1, as argued recently by Pajak and Levy (2014). Here we need to assume that listeners are able to decompose categories into features.

Pajak and Levy (2014) showed that selective attention to features works in the case of non-native contrast discrimination, depending on the phonemic or allophonic status of a given feature in the L1 as opposed to a situation where a given feature is absent from the L1 and does not positively influence foreign consonant perception. Here, we hypothesize that features familiar from the L1 play a role in the development of L2 perception and that in foreign language perception features from the L1, L2 and L3 can exert influence. The assumption is that listeners are able to decompose categories in such a way so as to isolate the features that are active in the languages they know and use them in speech perception. As vowels are composed of features or feature bundles and they obey some universal principles within each system, not all logically possible combinations of vowel features are attestable. This certainly makes testing more challenging and interpreting the results less straightforward.

Already, an individual familiar feature which is common in the L1 and the foreign language increases the likelihood of accurate perception of a foreign sound, so a familiar feature cluster should increase the likelihood of accurate perception of an L2/L3 or foreign sound exponentially. In the speech perception system reorganized by second language acquisition there is a hierarchy of features. As, cross-linguistically, tongue height distinctions are more numerous than tongue advancement distinctions, the tongue height feature is expected to predominate over the tongue advancement feature. As lip rounding is a feature typical for non-low back vowels and is marked in the case of front vowels, whereas non-low back unrounded vowels are even more marked, it is expected that the difficulty in perceiving the feature [+ rounded] or [- rounded] will depend on the tongue advancement of the examined non-native vowel.

2.6. The ideas for testing selective attention to features in second, third and non-native languages

The idea is that this book will examine the extent to which selective attention to features can play a role in non-native speech perception. Pajak and Levy (2014) proved selective attention to features to be at play in foreign consonant perception. Here, it will be tested whether selective attention to features influences the development of the L2, and whether the L3 and foreign language perception are influenced by the features present in the L1, L2 and L3.

Two series of experiments on vowel perception were carried out to examine the interplay between categories and features in speech perception and to allow for construction of a hierarchy of features. The first series of experiments comprised a longitudinal study of English vowel perception by Polish advanced learners of English. The other two studies tested Dutch and Turkish and Dutch vowel perception by three groups of learners: L2 English, L3 French and L3 Dutch.

The experiments were designed to reveal which vowels are difficult to discriminate (supposedly because the subjects, L2, L3 or foreign language listeners[1], cannot attend to certain features vital for their discrimination), which vowels are assimilated to native categories and which vowels can be claimed to be forming separate categories. Assimilation to native categories is expressed by high categorization rates and high goodness ratings. Very

[1] In this book all the subjects learnt their non-native languages, i.e. L2 and L3, in a formal classroom context. The term *foreign* is used to designate a language they had never had experience with, not a language learnt in a formal classroom context (cf. Krashen 1981 and Ellis 1984).

high goodness ratings should signal that a non-native sound is treated as a good exemplar of the native category. Relatively high goodness ratings may suggest that a native category is expanded to accommodate a non-native sound. In the experiment testing the perceived similarity of Dutch vowels to Polish vowels, the subjects are expected to perceive the vowels they actually categorize as belonging to a new category as substantially different from native categories (cf. Fox, Flege and Munro 1995). Good discrimination is expected between vowels which resemble native categories. If a vowel is not similar to a native category, but uses a combination of features used in the L1 or L2, its discrimination from other vowels should be enhanced (as in the case of foreign language front or central rounded vowels). This can also suggest that a feature distinguishing the two vowels in question has been noticed and isolated.

The vowels differed along F1, F2, F3, and therefore vector length. The hypothesis about the familiar feature cluster increasing the likelihood of accurate perception of L2 sounds will be confirmed if, for example, learners of French and Dutch, which use front rounded vowels, better discriminate the Dutch front rounded vowels than learners of English. Learners of Dutch should be slightly more efficient, because Dutch uses both lip rounding for front vowels and exploits tongue height and advancement in the high front part of the vowel space more than French does. French has three front non-low unrounded vowels /i, e, ɛ/ and three front rounded vowels /y, ø, œ/. Dutch has four front non-low unrounded vowels /i, ɪ, e, ɛ/ and the /ɛɪ/ diphthong and four front rounded vowels /y, ʏ, ø, œ/ and many of the Dutch vowels have longer and shorter counterparts (there is no agreement as to their phonemic status). French has front rounded vowels, but it disperses them more along the vertical dimension and it does not have /ʏ/. Therefore, all the contrasts involving /ʏ/ are supposed to be relatively challenging for learners of French, because they lack the features connected with tongue position in this area of the vowel space and less challenging for learners of Dutch, who are exposed to those vowels regularly in their L3. According to the hypothesis, subjects learning French should be much better at discriminating front rounded vowels than subjects learning English, because English does not use lip rounding for front vowels at all. English and Polish use lip rounding as secondary, non-distinctive features for non-low back vowels. As the results obtained by Pajak and Levy (2014) showed stronger effects for distinctive features than for secondary features, we might expect some gradation in the results.

Chapter Three

Experiments on the perception of English vowels by Polish advanced learners in a formal setting: A longitudinal study

This chapter is devoted to a longitudinal study of English vowel perception by advanced Polish learners of English. As most of the studies in this area were conducted with subjects learning the L2 in a natural immersion setting, the characteristics of a formal instruction setting will first be discussed. The following section will present the mixed results of research focusing on the role of language experience in speech perception. Further, the Polish and English vocalic systems will be described. Then, previous research on English vowel perception and production by Polish learners will be presented. On the basis of these introductory subsections, research questions will be asked. The methods section will describe the stimuli prepared for this study, and the subjects and procedure for the four tasks used: English vowel assimilation to Polish vowels with goodness ratings, English vowel identification (in stage two only), discrimination of pairs of English vowels and (dis-)similarity ratings. The experiments were conducted twice. The first session took place in the beginning of the subjects' studies at the Faculty of English, at Adam Mickiewicz University in Poznań, prior to pronunciation training. Then the subjects received the usual intensive language and phonetic training at the Faculty of English and were re-tested after two semesters of the course, when they had covered the English vocalic system, to see to what extent L2 speech perception difficulties can be overcome as a result of training or whether there are areas of difficulty which are related to specific features/feature combinations. Subsequent sections will present the results of the four tasks in both stages and discuss them.

3.1. Non-native speech perception in a formal instruction setting

There have been very few perception studies in classroom foreign language acquisition (henceforth FLA). Best and Tyler (2007) argue that in these circumstances (1) the target language is not widely used, (2) it does not extend much outside the classroom, (3) the emphasis is often on formal instruction on lexical and grammatical information rather than on live conversation, (4) the source of L2 input is either L1-accented speech or, at best, speech by native L2 speakers using diverse L2 varieties, thus learners are confronted with incorrect or variable models of L2 phonetic details. The points above led Best and Tyler (2007: 19) to conclude that "FLA is a fairly impoverished context for L2 learning, and perceptual findings for FLA listeners should not be conflated with those for L2 listeners (SLA)." They admit, though, that FLA learners' performance can be compared to SLA learners, because, even though their backgrounds cannot be tightly controlled, they have exposure to the target language. Nevertheless, the assumption in the present study is that even "exposure primarily through formal instruction in a restricted setting, with little or unsystematic conversational experience with native speakers" (Best and Tyler 2007: 19, emphasis theirs) can result in successful acquisition which fulfills educational and professional requirements, and as such and also as a means of L2 learning for millions of learners, warrants the need to study the process in more detail.

Moreover, the subjects in the present study are rather uniform: their level of English is advanced, which is confirmed by their final high school exams; they had a similar amount of exposure during classroom time; they were exposed to similar English listening comprehension tasks as a part of their curriculum; and they were taught by teachers with similar pronunciation training (in Poland pronunciation courses for English majors are intensive: approximately 180 hours of practical pronunciation training and 60 hours of descriptive English phonetics followed by oral and written examinations, and they are treated seriously).

Previous perception studies among learners acquiring language in non-naturalistic settings have been scarce, but they provide certain evidence that L2 learners learning the L2 in a formal, institutional setting may achieve near native-like speech perception in the L2. Bongaerts (1999) studied how Dutch learners of French perceive and produce L2 sounds and found that certain learner characteristics and learning contexts may promote the overriding of the disadvantages connected with a late age of onset of acquisition and a non-naturalistic learning setting: high motivation, massive exposure to the L2, and intensive phonetic training in L2 perception and production.

Mokari and Werner (2016) studied the link between the perception and production of vowels by Azerbaijani learners of English in a formal instruction setting. The results revealed correspondence between the discrimination of English vowels and their production (which was assessed both through acoustic measurements and native speakers' judgments).

3.2. The role of language experience in non-native speech perception

Both the Perceptual Assimilation Model (Best 1995) and the Speech Learning Model (Flege 1995) state that adult listeners assimilate non-native sounds to native categories. The Speech Learning Model stresses that in order to establish a new category for an L2 sound, learners have to detect differences between the L1 and L2 sounds. The more experienced the L2 learners are, the more likely they are to discern the differences between the L1 and L2 sounds and create new categories for the L2 sounds.

Several studies investigated the role of experience in L2 vowel categorization. Flege (1991) tested categorization of English vowels by Spanish listeners experienced in English and by monolingual Spanish listeners. Subjects listened to English vowels and were asked to circle one of the five letters used to spell the vowel in Spanish (<i, e, a, o, u>) or to circle "none" if they judged the vowel not to be found in Spanish. Experienced listeners circled the "none" label more often than Spanish monolinguals did (48% vs. 18%). This was taken to suggest that learning an L2 heightens bilinguals' awareness of phonetic differences between the L1 and L2 vowels. Certain subjects consistently used the "none" label for the vowels in *bit* and *bat*, but the group data did not suggest that Spanish listeners treated any English vowels as *new*.

Flege, Bohn and Jang (1997) assessed the effect of English language experience on German, Spanish, Mandarin and Korean listeners' perception of English vowels. The non-native listeners were divided into two groups based on their length of residence in the US. They were asked to identify vowels in synthetic *beat – bit* (/iː - ɪ/) and *bat – bet* (/æ - e/) continua. The experienced non-native listeners outperformed inexperienced listeners. Moreover, accuracy varied as a function of the L1.

Cebrian (2002 and 2006) did not find strong support for the effect of experience in L2 vowel identification or categorization. Cebrian (2002) compared the effect of experience in the case of Catalan learners of English in a formal instruction setting and in the case of Catalan learners in a natural immersion setting (Catalans who moved to Ontario, Canada after puberty). Subjects were asked to identify, in a forced-choice task, the naturally produced English /iː, ɪ, eɪ/ and /e/. The two groups of subjects produced comparable results, except for the more monophthongal productions of /eɪ/, which were better identified by more experienced learners. This was interpreted as less experienced listeners' reliance on the strong vowel offglide (they were based in Barcelona), which is not necessarily realized in Canadian English. In the case of the monophthongal /eɪ/, exposure to Canadian English promoted higher accuracy in categorization. Identification accuracy for /iː/ and /ɪ/ did not differ as a function of experience. The subjects identified longer /iː/ and shorter /ɪ/ tokens more accurately, which was interpreted as reliance on duration.

Cebrian (2006) further examined English /iː, ɪ, eɪ/ and /e/ vowel perception by native Catalan listeners who were either inexperienced or experienced in English. In the first experiment, the inexperienced subjects were native speakers of Catalan living in Barcelona, with minimal knowledge of and exposure to English. The experienced group consisted of native Catalan speakers who had lived close to Toronto, Canada for an average of 24 years. No effect of experience on English vowel assimilation to Catalan vowels was found. On the other hand, identification of L1 Catalan vowels varied between the inexperienced and experienced English learners. This suggested that there was a bidirectional influence of L1 and L2 vowels, as predicted by the Speech Learning Model (Flege 1995), which hypothesized that L1 vowel categories may be modified to fit L2 vowel categories in a single phonological space.

Cebrian (2006) also examined the role of L2 experience in reliance on temporal and spectral features in an identification task with a synthetic continuum involving English /iː/, /ɪ/ and /e/ vowels. Two groups of Catalan learners of English and a group of native English listeners participated in the second experiment. One group of learners included native Catalans who had lived in Toronto for an average of 25 years, whereas the second group consisted of third and fourth year undergraduates in English, living in Barcelona, and having limited experience in English-speaking countries. The tested stimuli were obtained from the two-dimensional /iː/ – /ɪ/ – /e/ continuum including 11 spectral steps and four temporal steps. The results revealed that native English listeners consistently relied on spectral cues in vowel categorization, irrespective of duration. Catalan learners of

English from both groups, however, relied strongly on duration when categorizing /iː/ and /ɪ/.

In Cebrian's (2006) study, the effect of experience in English was only evident in Catalan vowel identification by non-learners as opposed to learners. No effect of experience was found in the case of English vowel categorization by advanced learners living in Canada vs. advanced learners living in Catalonia (which was in line with Cebrian 2002). Subjects living in Canada had been more exposed to English, used the L2 more, and the L1 less. The undergraduates in Catalonia, however, had metalinguistic knowledge, had undergone phonetic training and were highly motivated learners. As the author notes, experience might exert greater effect in the case of less advanced and less motivated learners.

Rallo Fabra and Romero (2012) examined how well native Catalan listeners majoring in English and divided according to proficiency into three groups could discriminate Catalan-English and English-English vowel contrasts. As for Catalan-English contrasts, the subjects discriminated well between /i - ɪ/, partially between /i - i/, /u - u/ and /a - ɑ/, but found /a - ʌ/, /a - æ/ and /ɛ - e/ challenging. In the case of English vowel contrasts, the participants were able to distinguish between /iː - ɪ/ and /uː - ʊ/, which was interpreted by the authors as an indication of the possibility of new category formation. Generally, the higher the proficiency of the subjects, the more accurate their responses were. It is worth noting that Rallo Fabra and Romero (2012) used a rather challenging task to test vowel discrimination and accordingly, the results varied as a function of proficiency. The stimuli were /sVt/ with the final stop cut out, to prevent reliance on the dentoalveolar released stop production of Catalan /t/ and alveolar unreleased stop realization of English /t/. This step meant that English checked vowels /ɪ, e, æ, ʌ, ʊ/ were tested in open syllables, which violates English phonotactics. The three stimuli in each trial were produced by three different speakers, encouraging the subjects to ignore subcategorical variation. Further, each vowel contrast was tested in 16 trials: eight different trials, in which the odd item could appear in any of the three positions and eight catch trials, where all the vowels came from one category. The total number of trials was 112 for Catalan-English contrasts and 64 for English-English contrasts.

Concluding, among the reported studies, there was evidence of the role of experience in L2 vowel perception (Flege, Bohn and Jang 1997, Cebrian 2006 in the case of synthetic stimuli, Rallo Fabra 2012), and mixed effects

(Flege 1991, Cebrian 2002). No differences in L2 vowel assimilation and reliance on duration by both more and less experienced subjects were found by Cebrian (2006), although there were effects on L1 vowel identification. Also, in Cebrian (2006), the experienced subjects were those living in Canada, whereas the inexperienced subjects were undergraduates with phonetic training, so actually these two groups had different kinds of language experience, and the results should be treated accordingly.

3.3. Study comparing L2 consonant assimilation to L1 categories with L2 consonant identification

Park and de Jong (2008) used monosyllables consisting of English stops or fricatives and /ɑː/ to estimate the extent to which L2 identifications represent reliance on L1 categories. The underlying assumption of the study was that L2 sound identification could either be due to reliance on an L1 category or due to the process of establishing a new L2 category. Their results suggested that L2 identification accuracy could be related to L2-to-L1 mapping, but only if the assimilation of an L2 sound to an L1 category receives high goodness ratings and if mapping of L2 sounds to L2 categories is modulated by the goodness-of-fit of L2 to L1 categories. L2 sounds which are considered similar to L1 categories have confusability that is based on the probability that the L2 sounds mapped onto the L1 segments are distinct from one another. L2 segments with low goodness ratings exhibited identification rates which were much higher than the predicted ones. Park and de Jong (2008) noted that this result complied with Flege's (1995) distinction between similar and new L2 sounds and tried to account for the fact that some of the seemingly new sounds were still perceived with lower accuracy rates, as if the new categories were not yet fully efficient.

3.4. Polish vs. English vocalic systems

English has 11 monophthongs in stressed positions: /iː, ɪ, e, æ, ʌ, ɑː, ɒ, ɔː, ʊ, uː, ɜː/ and schwa /ə/ used in unstressed positions (Wells 1962, 1982, 1984, Gimson 1964, 1980, 1984, Deterding 1997, Upton, Kretzschmar and Konopka 2003, Hawkins and Midgley 2005, Hannisdal 2007, Bjelaković 2017).

In present day English, /uː/ and /ʊ/ are centralized or even slightly more fronted than central and have little lip rounding (Hawkins and Midgley 2005, Bjelaković 2017). Cruttenden (2014) described the fronting of /uː/ to /ʉ/ as well-established and the unrounding of /uː/ and /ʊ/ as more recent innovations. Hawkins and Midgley's (2005) and Bjelaković's (2017) acoustic study confirmed higher F2 for /uː/ and /ʊ/. There is slight diphthongization of /iː/ and /uː/ (Collins and Mees 2013).

British English /æ/ has been undergoing lowering and centralization from a mid-low front position (Wells 1982, Upton et al. 2003, Fabricius 2007 Cruttenden 2014, Collins and Mees 2013, Hughes, Trudgill and Watt 2012, Bjelaković 2017).

It is also assumed that F1 for /e/ is relatively high (Hawkins and Midgley 2005, Bjelaković 2017).

Further, the temporal characteristics of English vowels are complex. The division of English vowels is a matter of terminological disputes. For example, House (1961) and Crystal and House (1988) called the American English vowels short and long. Halle and Stevens (1969) used the terms lax and tense, though as Ladefoged and Johnson (2011) noticed, in English in most cases tense vowels are defined as occurring in stressed open syllables, but with no specific phonetic correlates (Ladefoged and Johnson 2011), while Daunmu (2016) criticizes the division into lax and tense as problematic for any theory. Kurath and McDavid (1961), Moulton (1990) and Collins and Mees (1981 and 2013) use the terms *free* and *checked*, which denote differences in both duration and quality between the respective vowel sets. Nevertheless, there is agreement[1] that English short/lax/checked vowels /ɪ, e, æ, ʌ, ɒ, ʊ/, except for the schwa, and assuming longer duration for /æ/, require a following consonant, and that long/tense/free vowels /iː, ɑː, ɜː, uː/ and diphthongs do not require a following consonant and may end a word.

In addition to vowel inherent potential duration as determined for free and checked vowels, English vowels are shortened before fortis consonants (Peterson and Lehiste 1960, House 1961, Chen, 1970, Raphael 1972 and Giegerich 1992) and lengthened in stressed positions (Fry 1955, 1958).

---

[1] Certainly, there may be alternative subdivisions, especially worthwhile when they take into account monophthongization of SQUARE and PURE diphthongs, and diphthongization of /iː/ and /uː/, and so called r-vowels (see Lindsey 2012), but as he predicts the traditional symbols for RP vowels will be used here, with the disclaimer that the symbols are meant to correspond to present-day English vowel qualities.

In comparison to the rich and complex English vowel system, Polish has a simple six-vowel inventory: /i, ɨ, e, a, ɔ, u/ (Wierzchowska 1980, Dłuska 1981, Jassem 2003) with no distinctions in tenseness or duration (though some of the vowels have nasalized variants). Table 1 presents mean Polish formant frequencies in Hz on the basis of recordings of seven male speakers taken by Nimz (2016). These data seem to be the most recent and reliable ones available for Polish vowel acoustics.

Table 1. Mean Polish formant frequencies for seven male speakers presented in Nimz (2016)

Polish vowel	F1 (Hz)	F2 (Hz)
/i/	283	2134
/ɨ/	392	1675
/e/	548	1559
/a/	637	1251
/ɔ/	504	992
/u/	347	829

3.5. Previous studies on English vowel perception and production by Polish listeners

On the basis of the comparison of English and Polish vowel systems above, it is clear that Polish learners of English need to learn to perceive about twice as many vowel contrasts as they have in their L1. This means learning to perceive both duration and smaller formant differences (both in the F1/F2 plane and in F3 relations indicative of lip rounding) as phonologically contrastive. Both perception and production studies with Polish learners of English have documented a considerable array of challenging areas (in addition to textbooks with many, mostly impressionistic, descriptions of articulatory differences: Jassem 1972, 1987, Bałutowa 1965, Krzeszowski 1968, Reszkiewicz 1981, Szpyra-Kozłowska and Sobkowiak 2011, Sobkowiak 2004, Porzuczek, Rojczyk and Arabski 2016, Sawala, Szczegóła, Jankowski, Weckwerth 2016). First, studies related to individual sounds will be reported, and the section will be concluded by a report on two studies of perception of English vowels by Polish learners of English.

Production studies focusing on the quality of English vowels have reported numerous difficulties which Polish learners of English are faced-with. Lipińska (2017) conducted a longitudinal study on /iː - ɪ/ production at the beginning and towards the end of the academic year, and she showed how first year English majors, in the post-test, but not in the pre-test, separated the two English vowels from each other and from their Polish counterparts. Szpyra-Kozłowska (2016) studied English loanword adaptation into Polish, and in her perception experiment, naïve Polish listeners assimilated the English /ɨ/ in around 50% of the cases to the Polish /i/ and she concluded that assimilation of the English /ɪ/ to the Polish vowel is not solely based on acoustic proximity, but on phonotactic constraints and other factors. Bogacka (2004) examined the perception of English high vowels by Polish learners and found that there was heavy reliance on temporal cues for both /iː-ɪ/ and /uː-ʊ/ continua and very weak reliance on spectral cues.

Rojczyk (2010a) studied the production of English /ɪ/ and /e/ vowels and found that third-year English majors separated the English /ɪ/ from the Polish /i/ and /ɨ/, but not the English /e/ from the Polish /e/ – although it must be noticed here that the second result is not surprising in the light of the reports of higher F1 for the English /e/ (see Hawkins and Midgley 2005 and Bjelaković 2017).

Gonet, Szpyra-Kozłowska and Święciński (2010a) examined production of /æ/ by first-year and third-year English majors. Assuming that the English /æ/ is a mid-low front vowel, they checked how well and when it is dissimilated from the Polish front mid /e/ and open front-centralized /a/. They reported considerably better results for more advanced subjects and noted that difficulties were context-dependent: productions were more accurate in stressed syllables, before palato-alveolars, obstruents and nasals. In unstressed syllables, in pre-stress positions especially, even advanced subjects had difficulties producing the ashes before anterior coronals and liquids. Further, the Polish vowel substituted for the English /æ/ depended on the vowel in the English loanword functioning in Polish.

Rojczyk (2010b and 2011) studied the production and perception of English checked vowels /æ/ and /ʌ/ by Polish learners. He found that /æ/ and /ʌ/ are spectrally subsumed by the Polish /a/: in production as formant frequencies for the two categories overlapped and in perception as the stimuli from the two vowel categories were not consistently or accurately identified. Both in production and in perception subjects relied on duration. Increased duration in production signaled /æ/. In perception, stimuli with longer durations tended to be identified as /æ/ and with shorter durations as /ʌ/, irrespective of their spectral characteristics, although Polish uses does not use duration as a cue to vowel identity.

Weckwerth (2011) studied the production of three English vowels: /e, æ /and /ʌ/ by 106 Polish students of English. He found good separation between /e/ and /ʌ/, whereas /æ/ was superimposed on the two, and, unlike in many dialects of English, it was not the lowest vowel. To show inter-speaker variability, a system typology was proposed. For some subjects the /æ/ tokens were situated outside of the ellipses for /e/ and /ʌ/, for others the majority of the /æ/ tokens were produced within the ellipse for /e/ or for /ʌ/. There was also a ‘bimodal’ system, in which a similar number of the /æ/ tokens fell in the space for /e/ and for /ʌ/, with less than 30% which fell outside of the two ellipses.

Schwartz, Aperliński, Kaźmierski and Weckwerth (2016) and Schwartz, Kaźmierski, Weckwerth, Jekiel and Malarski (under review) conducted studies of the production of English front vowels by Polish advanced learners of English: university students and teachers. They investigated the Vowel Inherent Spectral Change (Morrison and Assmann 2013), i.e. the dynamic formant trajectories over the course of a vowel’s duration. Its role is minimal in Polish, but crucial in the English vowel system. They found that the more advanced subjects produced more dramatic patterns of formant movement, closer to those of native speakers. Schwartz, Aperliński, Jekiel and Malarski (2016) confirmed the same tendency in L2 speech perception, i.e. the more advanced the listeners were, the more they attended to dynamic cues in vowel perception.

Several studies were devoted to the quality and duration of schwa in Polish accented English. Bogacka [Balas], Schwartz, Zydorowicz, Połczyńska-Fiszer and Orzechowska (2006) examined the production and perception of schwa by Polish adolescent learners of English and found that, depending on the position in a word, it was produced as either /e/ or /ɨ/, and perceived mainly as /e/, but actually many other targets appeared in subjects’ responses. Gonet, Szpyra-Kozłowska and Święciński (2010b) examined the production of schwa by first-year, third-year and fifth-year English majors and reported 31% to 67% appropriate productions of schwa as a function of proficiency. Word-initial and word-final productions exhibited more errors than word-medial positions. In addition, many mispronunciations were spelling-based.

Porzuczek (2010), in a longitudinal study, observed that schwa in weak forms of the preposition *to* was rarely devoiced by Polish first- and second-year majors English in comparison to native speakers. With reference to durational cues, he pointed to increased durations of schwas produced by Polish learners of English, but also to considerable improvements between the two stages of the study.

Rojczyk (2013) used synthesized stimuli based on the noun and verb *record* with flat and constant f0 contour (a primary cue for stress recognition in Polish) and manipulated vowel quality and duration. Subjects were able to base their perceptions on vowel duration and quality in the absence of a cue which they would use in Polish.

Szpyra-Kozłowska, Stasiak and Święciński (2014) conducted a longitudinal study of allophonic distinctions: aspiration, pre-fortis clipping, dark and clear *l*, syllabic consonants and the lack of audible release in stops. Within this set, pre-fortis clipping was the most difficult process, with the lowest correctness scores, and no long-term improvement after pronunciation training.

Other studies also pointed to more general problems with the use of duration cues by Polish learners of English. Waniek-Klimczak (2005) studied temporal parameters, i.e. VOT, closure duration and vowel duration in L1 Polish L2 English immigrants in Wisconsin: both early bilinguals and late bilinguals. Vowel duration was inspected in short and long vowels, as a function of the following consonant voicing, stress condition, position within a unit and the tempo of speech. The differences between the durations of long and short vowels were much larger for native speakers of English than for both groups of Polish-English bilinguals. As for the use of cues to voicing of the following consonant, two tendencies were observed for all the three groups of subjects: longer vowels and shorter closures in the case of long vowels followed by lenis consonants; and shorter vowels and longer closures in the case of short vowels followed by fortis plosives. The proportions in the use of the cues, however, varied: native speakers used the temporal difference to the greatest extent and late bilinguals to the least extent. Early bilinguals relied more on vowel duration, whereas late bilinguals relied on closure duration. The duration of each cue was exaggerated in comparison to the native speakers' productions. Generally, the more extreme values the subjects were supposed to produce, the more difficulties they had. Primarily stressed vowels were longer than secondarily stressed and unstressed vowels. Concluding, producing appropriate differences between long and short vowels, and vowel shortening before fortis consonants proved to be challenging for Polish learners of English.

Porzuczek (1996) examined the perception of English monophthongs and diphthongs by Polish advanced learners of English. His subjects had completed a one-year course in English phonetics, including transcription. In the experiment, they were asked to transcribe monosyllabic words, both existent and nonce, in which English vowels were embedded between a voiced or voiceless bilabial plosive and voiced or voiceless alveolar plosives or fricatives, bilabial nasals, bilabial plosives or /pV/or /Vt/. Short vowels were tested only in closed syllables. The stimuli were presented randomly to the subjects, who were asked to transcribe the words. Only /ɪ/ and /aɪ/ were transcribed correctly at rates higher than 90%. Scores higher than 80% were recorded for /əʊ/, /ɔː/ (confused with /ɒ/ and /uː/), /eɪ/ and /ɔɪ/. The /iː/ vowel was transcribed correctly at a rate of 79.6% and confused with /ɪ/. Almost 70% correct responses were given in the case of /ʊ/ (confused with /uː/) and /e/ (which was confused with /æ, ə/ and /ɜː/). The following vowels scored above 60% correct answers: /ɒ/ (confused with /ɔː and /əʊ/), /ɑː/ (confused with /ʌ/ and /aʊ/), /ʌ/ (confused with /ɒ, ɑː, aʊ/ and /æ/). Between 60% and 50% correct responses were given in the case of /uː/ (confused with /ʊ/), /aʊ/ (confused with /ʌ, ɑː/ and /æ/), /ɜː/ (confused with /eə, əʊ, ɔː/ and /æ/), /æ/ (confused with /ʌ, ɑː, e/ and /əʊ/). The diphthongs /eə/ and /ʊə/ only scored around 20% correct responses. Porzuczek (2007) also notes that perception was influenced by phonotactic constraints and familiarity with words or the similarity of nonce words to real words. The author stresses that there were many cases in which short vowels were mistaken for long vowels and long vowels were mistaken for short vowels. Importantly, short vowels were easy to identify before fortis codas and long vowels before lenis codas. Generally, English vowel length differences coupled with the effects of vowel length modifications depending on the following consonant (lenis or fortis) were very challenging for Polish learners. Porzuczek (1999) claimed that the relationships between English and Polish vowels were crucial, but not the only reason for misidentifications of English vowels by Polish learners. The following factors were shown to influence vowel recognition: the number of English vowels corresponding to the Polish vowel, associations with the word in which the vowel was being tested, context, frequency of occurrence of the vowel in English and spelling. It would be difficult to specify all the factors affecting identification of the English vowels and their relative importance.

Balas (2018) reported the results of English vowel perception by 35 advanced learners of English in the following tasks: discrimination task, assimilation task and (dis-)similarity ratings task. The paper had three main aims. The first one was to test whether learners from a formal instruction setting also rely on assimilation types when distinguishing between English vowels, as the Perceptual Assimilation Model predicts. Secondly, the effects of consonantal context on vowel perception were also examined. Unlike in the case of lower-proficiency learners (Strange, Akahane-Yamada, Kubo, Trent and Nishi 2001, Levy and Strange 2008), the place of articulation of neighboring consonants did not influence vowel perception, except for phonotactically-based restriction on /ɪ/ in velar contexts and allophonic fronting of high back centralized vowels in alveolar contexts. Thirdly, Balas (2018) also tested whether discrimination results conformed to the Natural Referent Vowel framework (Polka and Bohn 2003 and 2011) principles, according to which a change in a series of more peripheral vowels to a less peripheral vowel is more difficult to detect than the reverse change. Such asymmetries in perception had been previously attested for native and non-native vowels at six months of age, but only for non-native contrasts at 12 months and in adulthood. The results obtained by Balas (2018) failed to comply with Natural Referent Vowel framework predictions, which were interpreted as the ability of advanced foreign/second language learners to perceive foreign language vowels in a mode similar to native speakers, without referring to more peripheral vowels as anchors.

Some 18 of the 35 subjects retook the tests after a year of intensive English and pronunciation training and the differences between their results prior to the course and after the course will be discussed in the present chapter. Also, here we shall not focus on testing Perceptual Assimilation Model or Natural Referent Vowel predictions or the effects of consonantal context on vowel perception. The emphasis in the present study will be on tracking the changes in perception between stage one, before the onset of intensive English and pronunciation training, and stage two, i.e. after the two-semester courses in English phonetics and phonology and the role of features such as vowel height, backness and rounding.

3.6. Duration cues

Several competing hypotheses related to the role of duration cues in L2 vowel perception have been proposed so far. Bohn's (1995) desensitization hypothesis claims that whenever listeners do not perceive spectral differences between a given vowel contrast, because previous phonetic experience has not sensitized them to do so in a given vowel space area, they will rely on duration differences. McAllister et al. (2002) argue to the contrary: if the L2 features are not used contrastively in the L1, they will be difficult to perceive for the L2 learners. They actually did use duration to test their hypothesis. The hypothesis on selective attention to features (Pajak and Levy 2014) extends the feature hypothesis to cover the use of features in different phonetic environments and to emphasize the facilitative effect of reusing the known features. Lengeris (2009) argues that there is nothing in the complicated interplay of spectral and temporal cross-linguistics relations that would suggest increased or decreased temporal acuity dependent on the L1 experience. He argues that L1 transfer and assimilation types govern the perception of long and short vowels, similarly to the way they govern the perception of spectral differences.

As English has twice as many monophthongs as Polish, there are many areas in the vowel space that English uses and Polish does not. Therefore, if Bohn's (1995) desensitization hypothesis is true, it could be expected that duration differences help Polish subjects categorize, discriminate and notice differences between many English contrasts involving duration distinctions. If, however, the feature hypothesis and selective attention to features hypothesis are correct, then duration will not play a special role in non-native perception by Polish listeners.

The previous studies reported above, however, reveal that the issue is rather complex. There were studies which pointed to reliance on duration cues: Bogacka (2004) for the /iː-ɪ/ contrast and Rojczyk (2010b and 2011) for the /æ-ʌ/ contrast. Porzuczek (2010) showed improvement in the weak form *to* as a function of experience, and Rojczyk (2013) for the distinction between the verb and noun *record*. Waniek-Klimczak (2005), however, revealed that although both duration differences between long and short vowels and pre-fortis clipping were present, they were significantly weaker for Polish learners of English than for native speakers. Szpyra-Kozłowska, Stasiak and Święciński (2014) showed that pre-fortis clipping was very challenging for Polish learners of English.

3.7. Research questions

The present experiment is based on a longitudinal study of the perception of 11 British English monophthongs by Polish advanced learners of English acquiring L2 in a formal classroom setting. The following aspects of perception were investigated: discrimination; assimilation of English vowels to Polish vowels with goodness ratings; assimilation of English vowels to English vowels (in stage two only); and (dis-)similarity ratings. These studies allowed the following research questions to be addressed:

1. How are English vowels assimilated to Polish vowels and how do these assimilations differ before and after pronunciation training?

Theories of second language speech perception, such as the Perceptual Assimilation Model (Best 1995) and the Speech Learning Model (Flege 1995), assume that second language sounds are assimilated to the closest native categories. The closest native categories are not based on phonological proximity, but need to be established on the basis of perception tests. English has twice as many vowel phonemes as Polish does, so it can be hypothesized that two or more English vowels will be assimilated to one Polish category, but probably with varying goodness ratings. Goodness ratings in stage one are supposed to be higher than in stage two, when subjects, thanks to more language experience and phonetic training, had the chance to realize and notice that there are discrepancies between English and Polish vowels. Previous research has shown that learning the L2 raises learners' awareness of the phonetic distance between the L2 and L1 sounds (Flege 1991). This is one of the basic premises of category formation in the Speech Learning Model – one needs to notice the differences between L1 and L2 sounds in order to establish a new category for an L2 sound.

2. How are English vowels categorized in terms of English vowels after pronunciation training, when subjects are already familiar with transcription symbols? How are categorizations to English vowels related to categorizations to Polish vowels?

Park and de Jong (2008) showed that consonant assimilation to L2 categories can be predicted on the basis of L1 assimilations with goodness ratings. They also confirmed that some L2 sounds can function as categorized to L1 sounds, while for others listeners need new categories (cf. Flege 1995). Mayr and Escudero (2010) showed that cross-language assimilation patterns largely predicted identification of L2 vowels in terms of L2 vowels. It is therefore hypothesized that there will be a clear relationship between assimilation of L2 vowels to L1 categories and categorization of L2 vowels in terms of L2 vowels in such a way that if two vowels were assimilated as a Single Category to L1, then they will be confused with each other in an L2 identification task. Moreover, this part of the study will also allow comparison of the present results with the results obtained by Porzuczek (1996).

3. What are the discrimination rates of pairs of English vowels? Does discrimination improve? If so, for which pairs?

The predictions of the Perceptual Assimilation Model (Best 1995) have previously been confirmed for non-native, unfamiliar vowels (Tyler et al. 2014, Faris, Best and Tyler 2016) and second language vowels acquired in a formal instruction setting (Balas 2018). It is hypothesized in the present study that changes in discrimination rates for pairs of vowels should be related to changes in assimilation patterns and goodness ratings for these vowels.

4. How do (dis-)similarity ratings change overtime and how are they related to assimilation and discrimination results?

This study examines discrimination rates typically obtained for the Perceptual Assimilation Model studies and (dis-)similarity ratings to check whether they change in parallel. It is hypothesized that the more similar the two vowels are judged to be, the less discrepancy there should be in goodness ratings, if the vowels are assimilated to a common L1 category, and the lower the discrimination rate should be.

3.8. Method

With a view to answering the research questions posed above, a series of tasks was carried out to test the perception of English vowels by Polish advanced learners of English before and after two-semester long intensive language and pronunciation training for first-year English majors at Adam Mickiewicz University in Poznań. This section presents the stimuli used in the experiments, the subjects and procedure, which included the following tasks: discrimination of English vowels, assimilation of English vowels to Polish vowels, English vowel identification, and (dis-)similarity ratings.

3.8.1. Stimuli

The stimuli were the same as the stimuli used in Balas (2018). They were recorded by a male adult native speaker of Standard Southern British English in an anechoic chamber in the Center for Speech and Language Processing at Adam Mickiewicz University in Poznań. The speaker was instructed to read a list of carrier sentences containing nonsense[2] words to elicit production of 11 British English vowels /iː, ɪ, e, æ, ɜː, ʌ, ɑː, ɒ, ɔː, uː, ʊ/ in three homorganic CVC structures where lenis consonants /b d g/ were used. The vowel stimuli were both elicited and then presented to the listeners in closed syllables: bilabial /bVb/, alveolar /dVd/ and velar /gVg/. Closed syllables seemed to be the legitimate choice for investigating natural English vowel perception[3], as English checked vowels cannot occur in open syllables (Hammond 2003). Open syllables with checked vowels, which violate English phonotactics, should be avoided in natural speech perception experiments. Using syllable-final lenis consonants meant eliciting relatively long vowels (Rapheal 1975) and made the vowels more conspicuous than if they had appeared before fortis consonants. The speaker read the target sentences from PowerPoint slides presented to him in a booth. Two sentences appeared on each slide. The first sentence comprised real words which contained vowels which were identical to the vowels in the desired nonsense word in the second sentence below: In fork and Yorker we have /ɔː/. In borb and borber we have /ɔː/. The word which was cut out and later used for presentation in perception experiments was the monosyllabic word containing the vowel in question from the second sentence. In the example sentence, that would be the word *borb*. In each list containing the 11 vowels in a given context, the order of sentences with a given

---

[2] Three of these accidentally happen to be real words: *did, dad* and *bob,* but the author has not noticed any effect of familiarity on vowel categorization, goodness ratings, discrimination or (dis-)similarity ratings.

[3] At this stage it is difficult to predict whether or to what extent investigating the English vowels in isolation would influence the results for Polish learners of English, i.e. we do not know whether L2/FL learners learn to adhere to L2/FL's phonotactic restrictions in perception. This issue requires a separate study. The present author is only aware of preliminary evidence in favor of native phonotactics influencing the perception of L2 consonants (Bundgaard-Nielsen et al. 2016).

vowel was randomized. The speaker read four blocks of each list with a randomized order of vowels in each consonantal context. A Toshiba laptop computer and an Edirol UA-25 USB audio interface were used to make the recordings. The experimenter used English to communicate with the speaker and monitor the recording through headphones. During the recording session, the stimuli were digitized directly as computer files using Praat software (Boersma and Weenink 2015), with a sample rate of 22,050 Hz, 16-bit resolution, and a mono channel.

The speaker recorded four blocks in each consonantal context. The first block in each context was treated as a warm-up. The digital files from the second, third and fourth blocks containing the full sentences (e.g. in borb and borber we have /ɔː/.) were edited to isolate the nonsense monosyllables (e.g. borb). The tokens were selected in such a way that they matched in terms of pitch and average formant values. Then the selected tokens were examined by a monolingual British English speaker, who was able to identify all of them accurately and with reported ease. These tokens were used in perception tests. Figure 1 presents the average spectral characteristics of the 11 Southern Standard British English vowels produced by the speaker in the three contexts (bilabial, alveolar and velar) and chosen for the experiments in study one. F1 and F2 values were measured in Hertz at temporal midpoint of the CVC syllables. Table 2 gives mean formant values and vowel duration in the stimuli used in the experiments.

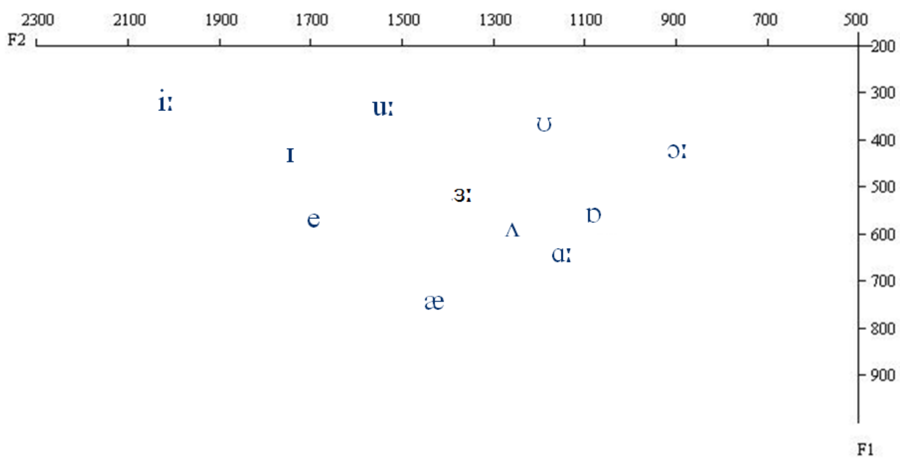

Figure 1. A plot of the vowel stimuli used in study one

Table 2. Mean vowel duration and formant values of the stimuli in study one

English vowel	Duration (msec.)	F1 (Hz)	F2 (Hz)
/iː/	224	311	2202
/ɪ/	117	430	1743
/e/	147	573	1691
/ɜː/	244	511	1376
/æ/	190	728	1415
/ʌ/	131	592	1264
/ɑː/	267	629	1159
/ɒ/	139	556	1057
/ɔː/	237	411	889
/ʊ/	110	366	1197
/uː/	207	322	1653

Comparing the formant values presented for Polish vowels (Table 1 in section 3.4.) with the vowel measurements presented here, we can assume that the Polish /i, ɨ/ and /e/ vowels are just slightly higher and slightly more centralized than their English counterparts /iː, ɪ/ and /e/. The Polish /a/ is lower than the English /ʌ/ and /ɑː/, but much higher than the English /æ/. The Polish /ɔ/ is closer to the English /ɒ/ than /ɔː/. The English /uː/ and /ʊ/ are lower and more centralized than the Polish /u/. Additionally, there are differences in length, lip rounding, diphthongization and formant dynamics.

3.8.2. Subjects

A total of 22 students, mean age 19 years 10 months (as reported at stage 1), all first-year English majors at the Adam Mickiewicz University in Poznań, Poland, took part in the experiment. They claimed not to have any hearing disorders. According to the Common European Framework of Reference for Languages, their knowledge of English was at the B2 level. This means they were upper intermediate learners who could understand the main ideas of a complex text or interact with a degree of fluency with native speakers. The tests in stage one were carried out at the beginning of the subjects' first academic year. The tests in stage two were carried out at the end of that academic year, after eight months of instruction, which included English pronunciation and English phonetics and phonology. All the examined subjects were enrolled in British English groups. They were recruited from four different instruction groups, so although their curricula were to a great degree similar, they received instruction from various teachers throughout the year. Two participants reported knowing another foreign language at a level higher than B1. B1-level learners can understand the main points of utterances on familiar matters and can talk about familiar matters. The majority of the participants had a basic knowledge of German, French or Spanish at A1/A2 levels. The responses from subjects admitted to the first year of studies for the second time were discarded, as the influence of instruction in the first year is a major factor possibly influencing the development of perception between stage one and stage two.

3.8.3. Procedure

A longitudinal perception study was carried out among first year majors in English. The first session was carried out in the first month of the subjects' first academic year. The second session was carried out after eight months of instruction in the subjects' first year of studies. In stage one the three perception tests were chosen and designed to evaluate the perception of English vowels by Polish listeners without the need to resort to phonetic symbols or orthographic labels. In stage two, the three tests were repeated and a fourth task, which assumed familiarity with phonetic symbols for English vowels, was added. Following previous studies testing Perceptual Assimilation Model's predictions for both consonants, e.g. Best and Strange (1992) and Best, McRoberts and Sithole (1988) and vowels (Tyler et al. 2014), the participants completed a discrimination task and a category identification task of English vowels in terms of Polish vowels with a goodness rating. Additionally, participants completed a (dis-)similarity rating task. The fourth task, carried out only in stage two, was a category identification task of English vowels in terms of English vowel labels supplemented with a goodness rating.

For the first test, participants completed a categorical discrimination oddity test. Ten contrasts were examined: /iː - ɪ/, /e - ɪ/, /e - æ/, /e - ɜː/, /e - ʌ/, /æ - ʌ/, /ʌ - ɒ/, /ɑː - ʌ/, /ɔː - ɒ/ and /uː - ʊ/. The vowel contrasts were chosen with aim of examining English vowels that are close to one another in the F1/F2 vowel space and were impressionistically reported to be challenging for Polish learners in various pedagogical resources. Each of the 10 vowel pairs was presented in a triad, where A was a stimulus from one English vowel category and B was a stimulus from another vowel category. The six triad combinations, i.e. AAB, ABA, ABB, BBA, BAB and BAA were presented randomly during the test. All the triad combinations were assigned one contrast each, with the exception of the BAA and BBA types, which were assigned two contrasts. Each contrast was presented to a given listener twice in two consonantal contexts (either bilabial, alveolar or velar) and in two trial types (AAB, ABB, BAA, BBA, ABA and BAB) (see Appendix 1), so that 40 responses were elicited from each subject. This step, which did not require all the contrasts in all consonantal contexts in all triad combinations, was taken to ensure that the task was not too long to prevent the other two tasks in stage 1 and three tasks in stage 2 from being performed during the same session.

To increase the focus on the phonetic category identity rather than the sheer physical identity of the stimuli, a categorical discrimination procedure was employed. In a triad, one stimulus was always the odd one, but the other two, while representing the same L2 phonological category, were never physically identical: $A_1A_2B_1$, $A_1B_1B_2$, etc. Such a procedure was designed to prevent listeners from making simple acoustic identity judgments. Such a paradigm tests only those acoustic differences that influence category identity, and avoids the dilemma of response notation in the form of phonetic symbols. Additionally, the interstimulus interval was considerably long to encourage phonological processing (ISI =1 s, and the intertrial interval = 6 s).

After the discrimination tests, participants performed an assimilation task with goodness ratings. They matched the English auditory stimulus with an orthographic Polish vowel label and rated the vowel's goodness in terms of its similarity to the chosen Polish vowel on a scale from 1 to 7, where 1 meant *barely matching the Polish vowel* and 7 meant *well matching the Polish vowel*. Six Polish orthographic vowel symbols: *i, y, e, a, o* and *u* acted as labels in the assimilation task. As Polish vowel orthography is transparent, using orthographic labels was judged to be clear to the subjects, who had not been familiarized with IPA vowel symbols prior to stage one of the experiment. In this task, each subject listened to each of the 11 English vowels in each consonantal context (bilabial, alveolar and velar). From each participant, 33 responses were elicited.

Only in stage two was a vowel identification task performed. This time subjects listened to the same set of 33 English stimuli (11 vowels in three consonantal contexts) and identified the auditory stimuli in terms of English vowel labels arranged in the shape of the vowel chart (see Figure 2). In this task, subjects were not asked to goodness-rate the English vowels in terms of English vowels, as it was assumed that once one identifies an English vowel in terms of an English vowel label, one does not need to rate whether it barely fits the category or matches it well.

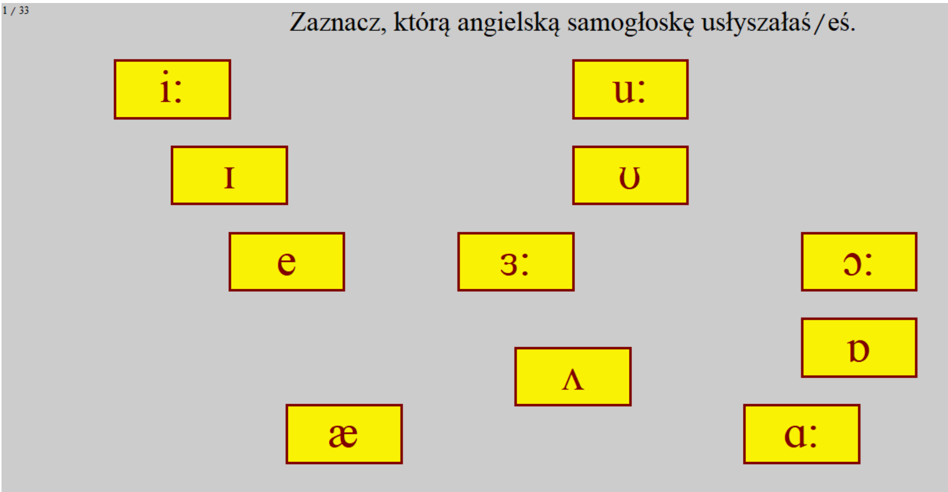

Figure 2. Screenshot of the screen presented to subjects in the identification task in stage two

Subjects also rated 12 pairings of Polish vowels with the 11 English vowels for their perceived (dis-)similarity: /iː - ɪ/, /e - ɪ/, /e - æ/, /e - ɜː/, /e - ʌ/, /æ - ʌ/, /ʌ - ɒ/, /ɑː - ʌ/, /ɔː - ɒ/, /uː - ʊ/, /ʊ - ɜː/ and /ɒ - ɑː/. Each pair was tested in three consonantal conditions (bilabial, alveolar and velar), yielding 36 vowel pairs to rate for each speaker. The stimuli were presented randomly. The participants were asked to indicate how dissimilar the vowels in the presented nonce words were, using a scale from 1 (barely similar) to 7 (really similar), and to make a guess if in doubt.

3.9. Results

Usually in the studies testing the Perceptual Assimilation Model, subjects complete discrimination tasks first and then proceed to categorization tasks. In the series of experiments reported on here, the (dis-)similarity task was completed as the last task. In the results section, however, the order is different. First, categorization results are presented, as assimilation types are needed to report discrimination results. (Dis-)similarity ratings will be presented in the final results and discussion sections.

3.9.1. English vowel assimilation to Polish vowel categories

The mean percentage of English vowel assimilations to Polish vowel orthographical labels are presented in Table 3, along with the mean category goodness ratings for these assimilations. These values were obtained by averaging all participants' ratings for a given stimulus in the case of a given Polish vowel label. An English vowel was deemed to be categorized if the same Polish vowel category was selected to represent it in more than 70% of the cases (following Antoniou et al. 2012, Bundgaard-Nielsen et al. 2011 and Tyler et al. 2014). A slightly different pattern of results for some of the vowels would have been obtained if the 50% criterion had been used.

Table 3. Mean percent assimilation and goodness rating (in parentheses) of English stimuli to Polish vowels

English vowel stimuli	Stage	Polish vowel orthographical labels					
		i	y (/ɨ/)	e	a	o	u
/iː/	one	98.15 (4.9)	1.85 (2)	–	–	–	–
	two	100 (4.5)	–	–	–	–	–
/ɪ/	one	46.3 (4.52)	50 (4.4)	3.7 (4)	–	–	–
	two	41.18 (4.3)	58.82 (4)	–	–	–	–
/e/	one	–	–	98.18 (4.7)	1.82 (2)	–	–
	two	–	–	100 (4.8)	–	–	–
/æ/	one	–	–	–	100 (4.7)	–	–
	two	–	–	1.96 (2)	98.04 (4.3)	–	–

/ʌ/	one	–	1.85 (2)	1.85 (4)	79.63 (4.5)	16.67 (3.8)	–
	two	–	–	1.96 (2)	76.47 (4)	21.57 (3.5)	–
/ɜː/	one	–	20.75 (3.9)	58.49 (3.6)	20.75 (3.3)	–	–
	two	–	15.69 (3.9)	66.67 (3.3)	15.69 (3)	1.96 (2)	–
/ɑː/	one	–	–	3.7 (1)	87.04 (4)	9.26 (3.8)	–
	two	–	1.96 (5)	1.96 (2)	90.2 (3.7)	5.88 (3.3)	–
/ɒ/	one	–	–	–	–	100 (4.8)	–
	two	–	–	–	–	100 (5)	–
/ɔː/	one	–	–	–	–	92.59 (4)	7.41 (3.5)
	two	–	–	–	–	88.24 (3.7)	11.76 (3.5)
/ʊ/	one	–	20.37 (3.5)	1.85 (1)	–	–	77.78 (4.4)
	two	–	13.73 (3.4)	–	–	–	86.27 (4.2)
/uː/	one	–	3.7 (5.5)	–	–	–	96.3 (4)
	two	–	–	–	–	–	100 (3.8)

T-tests were used to examine whether the changes in assimilation percentages between stage one and stage two were significant. At the significance level $\alpha = 0.05$ only the assimilations of /ʊ/ to the Polish /u/ changed significantly ($t = 1.761$, p-value = 0.0481). At the significance level $\alpha = 0.1$ the assimilations of the English /ɪ/ to Polish /ɨ/ ($t = 1.567$, p-value = 0.068) and of the English /uː/ to the Polish /u/ ($t = 1.458$, p-value = 0.082) also changed significantly.

T-tests were also used to examine whether, for a given English vowel, ratings of the goodness of fit to a Polish vowel were considerably lower in stage two in comparison to stage one. At the significance level $\alpha = 0.05$ only /æ/ and /ʌ/ were considered to resemble the Polish /a/ less in stage two than in stage one (t = -1.78025, p-value = 0.046457 for /æ/; and t = -2.08082, p-value = 0.010579 for /ʌ/). When the significance level was α = 0.1, goodness ratings were significantly lower in stage two, also for /ɑː/ (t = -1.59, p-value = 0.07) and /ɪ/ (t = -0.46, p-value = 0.09).

With reference to research question number one, the above results revealed which English vowels were assimilated to which Polish targets, and that assimilations to Polish vowels changed significantly only in three cases: /ʊ/, /ɪ/ and /uː/ and that goodness ratings were significantly lower in stage two for /æ/, /ʌ/ and /ɑː/.

3.9.2. Identification of English vowels

In stage two, when the subjects had already been made familiar with the phonetic symbols of English vowels, they were asked to categorize English vowels in terms of English vowels. On the screen they saw and clicked on the phonetic symbols of English vowels (see Figure 2, section 3.8.3).

Table 4. Mean percent identification of English vowel stimuli

Stimulus	English vowel labels										
	/iː/	/ɪ/	/e/	/ɜː/	/æ/	/ʌ/	/ɑː/	/ɒ/	/ɔː/	/ʊ/	/uː/
/iː/	**97**	3	–	–	–	–	–	–	–	–	–
/ɪ/	12.1	**83.3**	3	1.5	–	–	–	–	–	–	–
/e/	–	–	**90.9**	4.6	4.6	–	–	–	–	–	–
/ɜː/	–	4.6	1.5	**80.3**	4.6	–	7.6	–	–	1.5	–
/æ/	–	–	–	–	**51.5**	22.7	25.8	–	–	–	–
/ʌ/	–	–	–	1.5	12.1	**68.2**	4.6	10.6	1.5	1.5	–
/ɑː/	–	–	–	3	6.1	1.5	**81.8**	–	7.6	–	–
/ɒ/	–	–	–	–	–	–	–	**92.4**	7.6	–	–
/ɔː/	–	–	–	1.5	–	–	–	–	**89.4**	1.5	7.6
/ʊ/	–	13.6	–	3	–	–	–	–	–	**78.8**	4.6
/uː/	–	3	–	–	–	–	–	–	–	7.6	**89.4**

The mean percentages of English vowel label selections for English vowel stimuli representing a particular category are presented in Table 4. Boldfaced values indicate the most frequently chosen identification response per target. There were no category goodness ratings for these selections. An English vowel was deemed to be categorized in terms of an English vowel if the same English vowel category was chosen to represent it in more than 70% of the cases (following Antoniou et al. 2012, Bundgaard-Nielsen et al. 2011 and Tyler et al. 2014).

Research question number two asked how identification results were related to assimilation to Polish vowels in stage two. When we compare the results of the two tasks, we observe clear relationships. Vowels /iː/, /e/, /ɒ/ and /uː/ which were assigned to a given Polish category in more than 90% of the cases and received high goodness ratings (higher than the competitors) were identified with ease in the present task. In the case of Category Goodness assimilation types /ɒ - ɔː/ and /uː - ʊ/ we may observe that they were well identified in the present task. In the case of Categorized – Uncategorized partially overlapping assimilation types /iː/-/ɪ/ and /e - ɜː/ we observe that uncategorized vowels were identified in terms of English vowels with lower correctness scores. Vowels /æ/, /ʌ/ and /ɑː/ which were all assimilated to one Polish category /a/ with varied goodness ratings were identified with lower correctness scores here, and were, in fact, inversely proportional to the goodness ratings in the assimilation to Polish categories task. Concluding, research question number two regarding the relationship between the results of assimilation in terms of Polish vowels and identification of English vowels is answered positively, because the latter is dependent on PAM assimilation types.

3.9.3. Discrimination results

Research question number three asked about the rates of discrimination of pairs of English vowels and whether discrimination improves and if so, for which pairs of vowels. The overall mean percentage of correct discrimination responses for each contrast in stage one and stage two are presented in Table 5. The results are presented in decreasing order based on discrimination rates in stage one. The discrimination results for various contrasts varied, ranging in stage one from excellent discrimination for the /iː - ɪ/ contrast (98.86% correct responses) and five other contrasts (above 90% for /e - æ/, /e - ʌ/, /e - ɪ/, /ʌ - ɒ/, /ɔː - ɒ/, although the discrimination rates fell slightly below 90% correct for /e-ɪ/ and /ɔː-ɒ/), to very good discrimination for and /e - ɜː/ (84.09%), and fairly poor discrimination (below 80%) for /æ - ʌ/, /ʌ - ɑː/ and /uː - ʊ/ contrasts. The comparison of stage one and stage two discrimination rates did not reveal any statistically significant differences at the significance level of 0.05. At the significance level of 0.1, the difference in the discrimination rates for /uː - ʊ/ is significant. In addition to discrimination rates, Table 5 presents the calculations.

Table 5. Mean percent correct discrimination scores for the English vowel contrasts, and t-tests comparing the differences in the results between the two stages

Vowel contrast	Assimilation type in stage two	Percentage of correct discrimination in stage one	Percentage of correct discrimination in stage two	t-test	p-value
/iː - ɪ/	Categorized-Uncategorized, partially overlapping	98.86	96.59	-1.42241	0.158482
/e - æ/	Two Category	95.45	97.73	0.814937	0.417334
/e - ʌ/	Two Category	94.32	97.73	1.13576	0.259176
/e - ɪ/	Categorized – Uncategorized, non-overlapping	94.32	87.5	-1.61825	0.109229
/ʌ - ɒ/	Two Category	92.05	90.91	-0.29995	0.764932
/ɔː - ɒ/	Category Goodness	90.91	86.36	-1.06992	0.287613
/e - ɜː/	Categorized – Uncategorized, partially overlapping	84.09	88.64	1	0.320084
/æ - ʌ/	Single Category	78.41	71.59	-1.09671	0.275796
/ɑː - ʌ/	Single Category	76.14	82.95	1.13576	0.259176
/uː - ʊ/	Single Category	72.73	82.95	1.904433	0.060158

Assimilation types presented in Table 4 are based on calculations for the whole group, not on individual assimilation types checked against discrimination rates for a given subject (cf. Tyler et. al. 2014 and Balas 2018). Nevertheless, these calculations confirm the Perceptual Assimilation Model's predictions concerning discrimination rates as related to assimilation types. Two Category and Categorized – Uncategorized assimilation types in which either the two nonnative phones are perceived as acceptable exemplars of native phonemes or one is perceived as an acceptable exemplar and another as something that is not an exemplar of the same phoneme, should be characterized by very good to excellent discrimination. Tyler et al. (2014) called such assimilations Cross Boundary, because they cross a phonological boundary between two native phonemes and are both expected to be at least very well discriminated. Weaker correct discrimination rates are expected in the case of Category Goodness assimilations, because both nonnative phones are assimilated to a single native phoneme, nevertheless they do differ in goodness of fit to that phoneme. Sounds from Single Category assimilations, in which two nonnative phones are perceived as equally good or bad exemplars of a native category, are expected to be poorly discriminated. The results presented in Table 4 confirm these expectations, namely that vowels from Cross Boundary assimilations are the best discriminated contrasts, followed by Category Goodness and followed by Single Category assimilations. This finding gives a positive answer to research question number 3: discrimination rates depended on assimilation types, as predicted by the Perceptual Assimilation Model. Moreover, there was a significant improvement in discrimination only in the case of the /uː-ʊ/ contrast, where goodness ratings were low in stage one and fell lower in stage two, as if listeners noticed differences between the two English vowels and the Polish counterpart.

3.9.4. Results of dissimilarity ratings

In this experiment, participants rated the (dis-)similarity of two English vowels using a seven-point Likert scale. Twelve contrasts were tested: /iː - ɪ/, /e - ɪ/, /e - æ/, /e - ɜː/, /e - ʌ/, /æ - ʌ/, /ʌ - ɒ/, /ɑː - ʌ/, /ɔː - ɒ/, /uː - ʊ/, /ʊ - ɜː/, /ɒ - ɑː/. Three consonantal conditions, in which each contrast was tested, yielded 36 trials per subject. Upon hearing the two vowels embedded in two syllables with the same consonantal context, either /bVb/, /dVd/ or

/gVg/ and subjects were asked to indicate on a seven-point scale whether they thought the two vowels were not similar (1) or very similar (7).

The results of the (dis-)similarity rating task (see Table 6) are related to vowel discrimination results (see Table 5). Low vowel pairs were considered to be the most similar in stage one: /æ - ʌ/ (4.04) and /ɑː - ʌ/ (3.98) and in stage two: /æ - ʌ/ (4.32) and /ɑː - ʌ/ (3.46). There was a significant change in the dissimilarity rating for the latter pair, implying that subjects perceived more difference between /ɑː/ and /ʌ/ in stage two. These contrasts between low vowels also happened to be among the two worst discriminated ones (72 through 83% correct discrimination with significant improvement for neither pair) and with relatively similar and high goodness rankings. The two high rounded vowels /uː - ʊ/ were also considered similar (4.54 in stage one and 4.07 in stage two, a statistically significant improvement), most likely because the two features [+high] and [+rounded] (weakly rounded as English /uː/ and /ʊ/ are) are associated with only one Polish /u/ category which subsumes any L2 sounds sharing these two features, and disregards the tongue advancement difference. These two English vowels also posed difficulties in the discrimination task (73% correct discrimination in stage one vs. 83% correct discrimination in stage two, a statistically significant improvement) and they both had medium goodness ratings. It can be concluded that the three most poorly discriminated vowel pairs /uː - ʊ/, /æ - ʌ/ and /ɑː - ʌ/ were also the ones deemed to be the most similar.

Table 6. Mean (dis-)similarity ratings of English vowel contrasts on a Likert scale (1 meant *not similar*, and 7 meant *very similar*), and t-tests comparing the differences in the results between the two stages

English vowel contrast	(Dis-) similarity ratings in stage one	(Dis-) similarity ratings in stage two	t	p-value	Significance
/æ - ʌ/	4.04	4.32	-0.06522	0.948233	Non-significant
/ɑː - ʌ/	3.98	3.46	2.45917	0.017041	Significant
/uː - ʊ/	4.54	4.07	2.228547	0.029876	Significant
/ʌ - ɒ/	3.63	3.04	2.124111	0.03809	Significant
/iː - ɪ/	3.37	3.11	1.123734	0.26592	Non-significant

/ɔː - ɒ/	3.33	3.33	0.362355	0.718451	Non-significant
/ɒ - ɑː/	3.05	2.63	1.72628	0.089811	Non-significant
/e - ɜː/	2.96	2.93	0.179528	0.858171	Non-significant
/e - ʌ/	2.35	2.51	-0.80247	0.425676	Non-significant
/ʊ - ɜː/	2.27	2.21	2.376912	0.0209	Significant
/e - ɪ/	2.02	2.14	-0.61777	0.53923	Non-significant
/e - æ/	1.93	2.16	2.16	0.23264	Non-significant

There was also a significant improvement in the recognition of dissimilarity between /ʌ/ and /ɒ/. This vowel pair, although perceived as rather similar, at least in stage one, was rather well discriminated (92% correct discrimination in stage one, 91% correct discrimination in stage two, a non-significant change). There is a difference in the tongue height and the amount of lip rounding between the two vowels, which might be responsible for relatively good discrimination although the vector length differencies relatively small. Vowels /iː - ɪ/ were also considered only somewhat dissimilar (3.37 in stage one and 3.11 in stage two, a non-significant change), although their discrimination rate was ceiling in both stages. Vowels /ɔː - ɒ/ were also placed in the middle of the similarity-dissimilarity scale, but they were not discriminated at ceiling rates (90.91 in stage one and 86.36 in stage two). The /ɒ - ɑː/ contrast was judged as somewhat dissimilar (3.05 in stage one and 2.63 in stage two, a non-significant decrease). Vowels /e - ɜː/ were evaluated consistently as rather dissimilar (2.96 and 2.93 in stage one and two respectively), although their discrimination rates were not excellent (84.09% and 88.64%). In fact, both results were fairly close to the middle of the scales used by subjects in both tasks. Slightly weaker discrimination than we could expect on the basis of dissimilarity rating could be ascribed to the major difference between these two vowels being in tongue advancement only (duration did not matter for discrimination of other pairs, either). Contrast /e - ʌ/ was judged to be dissimilar (2.35 and 2.51 in stage one and two respectively) and it was also discriminated excellently: 94.32% and 97.73%. Vowels /ʊ - ɜː/ were also rated as dissimilar (2.27 and 2.1 in stage one and two, respectively). The two most dissimilar contrasts /e - ɪ/ (2.02 and

2.14) and /e - æ/ (1.93 and 2.16 in stage one and two, respectively) had discrimination rates higher than 90% and they were all categorized into different categories with the English /e/ categorized as Polish /e/, /ɪ/ and /æ/ not falling into the /e/ category. Vowel /e/ and the two other vowels in the most dissimilar contrasts are differentiated by vowel height differences.

3.10. Discussion

This study was a longitudinal examination of English vowel perception by Polish learners of English. The purpose was to test how selective attention to features helps in the development of L2 vowel perception. The results are discussed with reference to the four research questions and the role of selective attention to features. Discrimination results are discussed first, followed by assimilation, identification and (dis-)similarity ratings results.

3.10.1. Discussion of discrimination results

Research question number three asked about discrimination rates and their improvement between stage one and stage two. Only in the case of Single Category assimilation of /uː/ and /ʊ/ to the Polish /u/ did we notice statistically significant improvement in discrimination rates. We may only speculate why discrimination rates for low vowel pairs, also assimilated as Single Categories, did not improve – it may be because of minimally higher goodness ratings. It would seem that the more subjects notice the difference between the L2 sounds and the L1 sounds, the more likely they are to use the discrepancy to discriminate between the two sounds in question.

No improvement in discrimination between other English vowels was observed, which is an unexpected result. It was assumed that there would be improvement after a whole year of intensive pronunciation and language training. Surely, if assimilation types between stage one and stage two did not change, according to the Perceptual Assimilation Model we would not expect differences in discrimination rates. Differences in discrimination rates could also stem from lower goodness ratings in stage two. It needs to be emphasized here that lower goodness scores were only recorded for the low vowels /æ, ʌ, ɑː/ and for /ɪ/. What changed in the case of /uː/ and /ʊ/ were assimilation scores to the Polish /u/. They also changed for /ɪ/, but the /iː - ɪ/ contrast was excellently discerned, including in stage one, probably because it resembles the Polish contrast /i - ɨ/.

3.10.2. Discussion of categorization results

Research question number one asked about assimilation patterns of English vowels in terms of Polish vowel labels and also about goodness ratings. Moreover, it asked how categorizations and goodness rating differed in stage one and stage two. When we look at responses to given stimuli, we generally observe that the fewer the responses to a given target, the lower the goodness rating. This suggests that the measures were reliable. The results are in line with those presented by Balas (2018)

An examination of assimilations and goodness ratings reveals some interesting patterns. Beginning with the /iː - ɪ/ contrast, it needs to be noticed that there is asymmetry in perception here. The English /iː/ was perceived almost exclusively as Polish /i/, with average goodness ratings of 4.9 in stage one and 4.5 in stage two on a seven-point Likert scale.

The English /ɪ/, however, was also perceived as Polish /i/ in 46.3% and 41.18% of the cases in stage one and stage two, respectively, and with goodness ratings of 4.52 and 4.3 points. Assimilations to /ɪ/ amounted to 50% and 58.82% with goodness ratings of 4.4 and 4 in stage one and two, respectively. These assimilation results are in line with those reported by Szpyra-Kozłowska (2016), who examined English /ɪ/ adaptations in loanwords in Polish. She found a similar pattern in an /ɪ/ categorization study, which led her to conclude that the English /ɪ/ is mostly nativized as Polish /i/ in line with nativization through production (LaCharité and Paradis 2005) as a substitution which is phonologically, but not phonetically or perceptually, motivated. If we analyze assimilations of English /ɪ/ as a function of a consonantal context, it turns out that /ɪ/ in a velar context was assimilated as the Polish /i/ (94.44%, goodness rating 4.65 in stage one, the remaining cases were assimilated to /e/, none to /ɨ/; and 72.22%, goodness rating 4.92 in stage two). These results need to be ascribed to Polish phonotactics, which does not allow for /ɨ/ after velars. Polish learners of English, upon hearing /gɪg/ "repair" the vowel in an illegal /CVC/ string and choose the nearest legal /CVC/ counterpart as a target in perception. It must be noted here that the perception of second language vowels is governed by co-occurrence restrictions on consonant-vowel combinations in the L1, in addition to L1 and L2 vowel characteristics. In a similar fashion, Bundgaard-

Nielsen et al. (2016) described this effect for English consonants when perceived in illegal Japanese /VCV/ strings by Japanese learners of English. The authors noticed that models of non-native and cross-language speech perception like PAM can predict perception success for non-native contrasts, but fail to account for the role of native phonotactics in non-native segmental perception. The present results (in addition to Balas 2018, where 80% of /ɪ/ instances in a velar context were perceived as /i/) provide yet another piece of evidence that perception is modified by phonotactics. It is worthy of note that we can also observe a statistically significant increase in assimilations to /ɨ/ from 50% to 59.26 % (at $\alpha = 0.1$, $t = 1.573$, p-value = 0.068). The change is caused by an increase in assimilations of English /ɪ/ in a velar context to Polish /ɨ/ – from 0 in stage one to 27.78% in stage two. This implies that gradual overcoming of native phonotactics in L2 vowel perception is underway.

English /e/ instances were perceived as the Polish /e/ in 98.18% and 100% of the cases in stage one and stage two, respectively, and with relatively high goodness rankings (4.72 and 4.86) – probably due to a small articulatory difference in the vowel height. This result sheds new light on the interpretation of the result obtained by Rojczyk (2010a) – the English vowel /e/ is indeed rather similar to the Polish counterpart.

At the assumed 70% categorization threshold, the English /ɜː/ is an uncategorized vowel in both stages. It was assimilated to the Polish /e/ in 58.49% of cases, and had a goodness rating of 3.58 in stage one, and a 66.67% assimilation rate as well as a goodness rating of 3.2 in stage two. For other vowels, goodness ratings were lower in stage two than in stage one, signaling that subjects were aware of and/or perceived more differences between English and Polish vowels. Here, the t-test shows that, in fact, in stage two subjects were more confident that the Polish /e/ was the right target for the English /e/ ($t = 0.048$, p-value = 0.076). Other targets for the English /e/ included /ɨ/ and /a/ (around 20% assimilation in stage one, and 15% in stage two). The acoustic difference in both F1 and F2 between the English /ɜː/ and the Polish /e/ was perceived as being more striking than just a difference in vowel height between the English /e/ and Polish /e/.

Subjects treated the central Polish /a/, the only low vowel in Polish, as the target for the three English low vowels. Naturally, assimilation percentages and goodness ratings of the sounds from the three categories varied slightly. In stage one, the three English vowels scored as follows: /æ/

– 100% assimilation to the Polish /a/, 4.74 points in goodness rating, /ʌ/ – 79.63%, 4.51 points in goodness rating, and /ɑː/ – 87.04% and 4.04 points in goodness rating. In stage two, the following values were recorded: /æ/ – 98.05%,4.32 points in goodness rating, /ʌ/ – 76.47%, 4 points in goodness rating, and /ɑː/ – 90.2% and 3.74 points in goodness rating. None of the vowels exhibited a significant increase in assimilation percentages. The vowel /ʌ/, however, had considerably lower goodness ratings in stage two than in stage one ($t = -2.08$, p-value = 0.01 at $\alpha = 0.05$) and this was similar for the vowel /ɑː/ – $t = -1.59$, p-value = 0.06, but this time at $\alpha = 0.1$. Therefore, it can be stipulated that at stage two subjects noticed that these vowels differed considerably from the Polish target. This is the first prerequisite for new category formation according to Flege (1995).

As for phonetic details of the stimuli, the formant values for the /æ/ in here were typical of the current British English centralized low /æ/ as reported by Hawkins and Midgley (2005), as opposed to the former front /æ/. As at both stages /æ/ has the highest percentage of assimilations to the Polish /a/, and the highest goodness ratings, subjects must have considered it to be the best perceptual counterpart of the Polish /a/. The present results are in line with Rojczyk (2010b, 2011), who found that /æ/ and /ʌ/ were spectrally subsumed by the Polish /a/, and that irrespective of their spectral characteristics, stimuli with longer durations were identified as /æ/ and with shorter durations as /ʌ/. In production, the English /æ/ is commonly substituted by the Polish vowels /e/ and /a/, based on a 'false friends', segmental and suprasegmental context (Gonet et al. 2010a). Production results obtained by Gonet et al. (2010a) also revealed that the Polish vowels /e/ and /a/ which were substituted for the English /æ/ were probably chosen on the basis of spelling.

It seems that the more retracted an low English vowel is, the worse the goodness ratings it received. Also, between 16-21% of English /ʌ/ cases were categorized as the Polish /o/, and as expected the goodness ratings were relatively lower – 3.7. These results also suggest that any feature which is different from the features present in an L1 category (i.e. more retraction of the tongue or longer duration/tenseness in /ɑː/ and higher tongue position for /ʌ/) yields lower goodness ratings. Generally, discrepancies in vowel height for other vowels were easier to spot and penalized more than those in the tongue advancement.

As for the English mid back vowels /ɒ/ and /ɔː/, the subjects seemed to judge /ɒ/ to be the best exemplar of the Polish /o/, as 100% of its instances were categorized as /o/, and the goodness ranking's average was high – 4.83 points in stage one and 4.98 in stage two. The English /ɔː/ was perceived as the Polish /ɔ/ in around 90% of the cases in both stages, with goodness ratings ranging from 3.98 to 3.67. The remaining responses, although below the chance level, pointed to the Polish /u/ as being an assimilation target for the English /ɔː/, with a goodness ranking of 3.5 points. If the height of the Polish vowel is intermediate between the two English vowels /ɒ/ and /ɔː/, the tenseness of /ɔː/ must contribute to the vowel being judged as a worse exemplar of the Polish /o/ than the English /ɒ/.

The English /uː/ was perceived as Polish /u/ in 96.3% in stage one, with an increase to 100% in stage two. The increase in assimilation to the Polish /u/ was statistically significant at $\alpha = 0.1$, $t = 1.46$ and p-value = 0.081569. Goodness ratings equaled 4.04 points in stage one and 3.8 in stage two, but the change was not statistically significant. In stage one, the only instances of /uː/ which were not assimilated to the Polish /u/ were the /uː/ stimuli in the alveolar contexts assimilated to the Polish /ɨ/ – 11.11% with a goodness rating of 5.5, so considerably better than for assimilations to /u/ in the same context, i.e. 3.88 (which is identical to the goodness ratings in the bilabial context and much worse than in the case of the velar context – 4.39). These details must be due to /uː/ fronting in the alveolar context and the retracted realizations of /uː/ in the velar context. Subjects gave low goodness ratings to fronted realizations in the alveolar context, central realizations in the bilabial context, and clearly preferred the back /uː/ in the velar context. The length of the vowel stimuli and the limited lip rounding still make such back realizations of /uː/ in the velar context not the ideal fit for the Polish /u/.

Fewer instances of /ʊ/ than of /u/ were assimilated to the Polish /u/ – 77.78% in stage one with an increase to 86.27% in stage two. The change was statistically significant at $\alpha = 0.05$ with $t = 1{,}76$ and p-value = 0.048. Goodness ratings decreased from 4.36 in stage one to 4.2 in stage two, which was not a significant drop. The remaining /ʊ/ instances were assimilated to the Polish /ɨ/ – 20.37% in stage one and 13.73% in stage two. If we look at consonantal contexts, we can notice that no instances of /ʊ/ in velar contexts were assimilated to /ɨ/, only 16.67%, which happens to be the chance level, of instances of /ʊ/ in the bilabial context, and 44.44% in the alveolar context (here the goodness ratings were also higher than for other targets) were assimilated to the Polish /ɨ/. The last case needs to be interpreted as a result of /ʊ/ fronting in alveolar contexts. To disentangle the relationships between vowel formant characteristics, contextual influences and the role of duration/context, a more detailed study involving the three contexts and manipulated duration steps would be needed.

The results of stage one in the present study do not considerably diverge from the results presented in Balas (2018), which is not surprising, because 18 of the 35 subjects' results in the other study were actually analyzed here. The main focus of this study, however, is the change in perception between stage one and two. When we analyze assimilation results, it seems that in stage two those assimilation targets which were most often chosen in stage one actually gather even more responses, in the vein of "winner takes all". The increased assimilation percentages were statistically significant in the case of /ʊ/ (at α = 0.05) and /ɪ/ and /uː/ (at α = 0.1). The results were ceiling in stage one for the following vowels: /iː, e, æ/ and /ɒ/, so they could not be increased. The changes were not statistically significant for /ɜː, ɑː, ɔː/ and /ʌ/.

In the light of phonetic information and comparative English-Polish phonetics and phonology, the change in the case of /ɪ/ was caused by fewer assimilations of /ɪ/ to Polish /iː/ in the velar context, where /ɨ/ is not allowed by Polish phonotactics. We can observe here a gradual process of restricting the application of a Polish phonotactic rule in L2 speech perception.

In the case of /uː/ and /ʊ/, the increase of assimilations to /u/ must have been caused by embracing the fronted realizations in alveolar contexts – in stage two subjects showed that they had begun to realize that the English categories for /uː/ and /ʊ/ are wider than for the Polish /u/, that they involve fronted variants and that these vowels are not as rounded as the Polish vowels.

Subjects in the present study complied with the requirements stipulated by Bongaerts (1990) for successful L2 perception: as English majors, they were highly motivated, had considerable exposure to the L2, and they underwent intensive phonetic training between stage one and stage two of the experiments. Both the Perceptual Assimilation Model (Best 1995) and the Speech Learning Model (Flege 1995) assume that L2 learners assimilate L2 sounds to L1 categories. With growing experience in the L2, they should start noticing more discrepancies between the L1 and L2 sounds, so that they have the motivation to establish new categories for the L2 sounds. Flege (1991) and Flege et al. (1997) showed that experienced learners outperformed inexperienced learners. Cebrian (2002) did not show any differences in L2 vowel categorization by immersion learners and learners from a formal education setting. Cebrian (2006) found no difference in English vowel assimilation to Catalan categories in the case of advanced learners (English majors in one group, immigrants to Canada in the other). He noted that there might be a bigger difference at lower levels of proficiency.

Bearing these mixed results in mind, it can be argued that the present study did manage to trace the change of perception in progress. In the case of three categories /ʊ, uː/ and /ɪ/ an increase in assimilation percentages for a given Polish target was observed. More importantly for new category formation (Flege 1995), in the case of five English vowels, /æ, ʌ ɑː/ and /ɪ/, a significant decrease in goodness ratings was found. For advanced learners of English, in spite of several ceiling results, we still observed changes in the case of five vowels. This implies that even for advanced learners, modifications in perception are attainable.

All three English low vowels /æ, ʌ/ and /ɑː/ were assimilated to the Polish /a/ with varying goodness ratings. The results for /æ/ were 100% and 4.7 in stage one, and 98.04 and 4.3 in stage two. The decrease in goodness ratings was significant. The English /ʌ/ was assimilated to Polish /a/ in stage one in 79.63% of the cases (goodness rating: 4.5) and 76.47% of the cases (goodness rating: 4). The decrease in goodness rating was also statistically significant. The contrast /æ-ʌ/ was assimilated as Single Category type. The English /ɑː/ was assimilated as Polish /a/ in stage one in 87.04% of the cases (goodness rating 4) and in stage two in 90.2% of the cases, goodness rating (3.7). The decrease in goodness ratings here was also statistically significant, but at $\alpha = 0.1$ ($t = -1.59$ and p-value = 0.06). The contrasts /æ-ɑː/ and /ʌ-æ/ were assimilated as Category Goodness types. It seems that one or both features differentiating /ɑː/ from the Polish /a/, either duration/tenseness or considerable retraction, must have been responsible for lower goodness ratings for /ɑː/ in comparison to /a/. The fact that all the three English low vowels were assimilated to /a/ with varying goodness ratings signals that subjects noticed differences at least between /ɑː/ and the other two more fronted English vowels. This was probably due to selective attention to features: listeners accepted all the three English vowels as counterparts of the Polish /a/, but marked that not all of them fitted the Polish category equally well.

On the basis of the goodness ratings for /ɜː, ɑː, ɔː/ and /uː/, which were lower than for checked vowels, it would seem that longer duration/tenseness was responsible for the results. Goodness ratings for /iː/, however, were generally high. In terms of articulatory characteristics, the

English /iː/ is closer to the Polish /i/ than other English free vowels are to their respective Polish counterparts. The English /ɜː/ is a central vowel, more retracted than the Polish /ɨ/ or /e/, and much higher than /a/. The English /ɑː/ is more retracted than the Polish /a/, British English /ɔː/ is much higher than the Polish counterpart, and the English /uː/ is central rather than back and has little lip rounding. So perhaps it is the coupling of two different features (duration/tenseness and spectral characteristics) that yields lowered goodness ratings.

Concluding, when we compare stage one to stage two, with reference to research question number one, the responses to the main targets increased in the case of /ʊ/, /uː/ and /ɪ/, as if "the winner takes all". Goodness ratings decreased in the case of /æ, ʌ, ɑː/ and /ɪ/, implying that subjects had realized/noticed that these vowels were different from their Polish counterparts. This step is necessary for establishing a new category. The English /e/ and /ɒ/ were considered to be similar to their Polish counterparts in both stages and in fact they are articulatorily rather similar. The English /ɜː/ and /ɔː/ were considered dissimilar from their Polish counterparts in both stages. When we analyze differences between goodness ratings for English vowels or when we look at decreased goodness ratings in stage two, we can easily relate them to articulatory and acoustic differences between Polish and English vowels. Analyzing the cases of vowels which are different in respect of both spectral characteristics and duration/tenseness, we observe that two different features between the English vowel and the Polish counterpart yielded lower goodness ratings, which supports the hypothesis of selective attention to features.

3.10.3. Discussion of identification results

In stage two, subjects were asked to identify English vowels using English vowel phonemic labels based on the IPA alphabet, which they were familiar with from their pronunciation and phonetics and phonology courses. Correct answers prevailed, but they spanned from 97% for /iː/, around 90% for /ɒ, e, ɔː/ and /uː/, around 80% for /ɪ, ɑː, ɜː/ and /ʊ/, 68.2% for /ʌ/ and as low as 51.5% for /æ/.

As predicted, those vowels which were assimilated as Single Category, namely /æ/ and /ʌ/, were identified with the lowest accuracy. Vowels in other contrasts were assimilated to a Polish vowel as Cross Boundary types, namely Categorized – Uncategorized and Two Category assimilation types, and they were better identified in the present task. Uncategorized vowels /ɪ/ and /ɜː/ were identified at correctness levels similar to /ɑː/ and /ʊ/, which were involved in Category Goodness assimilation types with medium goodness ratings differences (/uː/ – 3.8, /ʊ/ – 4.2, /ɑː/ – 3.7 and /ʌ/ – 4). The third example of Category Goodness assimilation was /ɒ - ɔː/ with a larger difference in goodness ratings: 4.8 and 3.7 respectively, and these two vowels were identified with high correctness rates, together with /e/ (where assimilation to the Polish /e/ did not show much overlap with other English vowels) and /uː/, which may not have high goodness rankings, but was unanimously assigned to the Polish /u/. The best identified vowel, /iː/, was involved in a Categorized – Uncategorized assimilation type, only partially overlapping with /ɪ/, but for phonotactic rather than acoustic reasons.

Generally, the vowels that were well identified, probably because they are conspicuous and salient, were the long/free vowels /iː/, /ɔː/ and /uː/, with the exception of /ɜː/ which has spectral characteristics that are new for L1 Polish listeners and /ɑː/ which is assimilated to Polish /a/ with two other English vowels. Also, vowels similar to Polish vowels (with high goodness ratings) and without strong competitors, i.e. /e/ and /ɒ/ were accurately identified. Vowel /ɪ/ is also rather similar to the Polish counterpart, but there are phonotactic restrictions on its occurrence (in Polish, /ɨ/ does not occur after velars), which have interfered here: /ɪ/ was correctly identified in 90.91% of cases in the bilabial context, 95.45% in the alveolar context and only in 63.64% in the velar context. Another fairly well identified vowel was /ʊ/ (78.8% correct identifications), which had higher goodness rating scores than its free/long counterpart /uː/. The results for low vowels would be another example of the tendency that free/long vowels are more accurately identified than their checked counterparts: the free and long /ɑː/ was more accurately identified. At the same time, the vowel /æ/ which is a checked vowel, but with long duration, was the worst identified vowel, though we may speculate that it was due to the two competing vowels, assimilated to the same Polish category as /æ/.

Addressing research question number two, we can conclude that identification accuracy of English vowels depended on their assimilation types to Polish vowels, that free/long vowels were easier to identify. Furthermore, in the case of Single Category and Category Goodness assimilations with relatively small differences in goodness ratings, the lower the goodness ratings for a given vowel were, the better it was identified.

Also, new tongue advancement contrasts were difficult: the three challenging low vowels /æ, ʌ/ and /ɑː/ are primarily distinguished by F2, and /ɜː/ and /ʊ/ represent high-mid and mid central vowels, whereas the only Polish central vowel is a low /a/. When the high central vowel was the long /uː/, its characteristic length made it more readily identifiable.

If a vowel was misidentified, the chosen response category was close to the target in terms of F1 and F2, but duration mediated categorization: /ɜː/ interpreted was as /ɑː/, /ʌ/ was interpreted as /æ/ (12.1%) or /ɒ/ (10.6%) and rarely as /ɑː/ (4.5), /ɑː/ was interpreted as a checked vowel, but with long duration – /æ/ (6%) or /ɔː/ (7.6%), but never as /ɒ/ or rarely as /ʌ/ (1.5%), /ɔː/ was interpreted as /uː/ (7.6%) much more often than as /ʊ/ (1.5%), which is closer to /ɔː/ in terms of F1.

In comparison to the transcription task results reported by Porzuczek (1996), the present results on accuracy in English vowel identification are generally higher, which may have to do with the lower cognitive load of the present task, where the subjects needed to choose one transcription symbol out of 11 monophthong symbols visible on the screen rather than transcribe the vowels, when the task involved both monophthongs and diphthongs.

3.10.4. Discussion of (dis-)similarity rating results

As Flege, Munro and Fox (1994) argue, L2 vowels should seem dissimilar in a crowded vowel space, if L2 learners are to establish new categories for them. The higher the perceived dissimilarity between L2 vowels, the more chances there are for establishing new phonetic categories.

Research question number four asked about the relationship between assimilation, discrimination rates and (dis-)similarity ratings. In the present study, similarly to Balas (2018), poorly discriminated pairs (/æ - ʌ/, /ʌ - ɑː/ and /uː - ʊ/) were the ones judged to be most similar. The fact that low vowels were rated as similar is yet another argument for the three vowels functioning as one Polish /a/ category with various degrees of goodness. The /uː-ʊ/ contrast was the worst discriminated contrast, the vowels were judged to be very similar and they were assimilated to a single Polish /u/ category. In both tests, however, the changes between stage one and stage two were statistically significant, which confirms the correlation between the two tasks. Two more contrasts (/ɒ - ɔː/ and /ʌ - ɒ/) deemed to be somewhat similar were discriminated at rates higher than

86% in both stages. Vowel contrasts considered to be dissimilar were the ones which were assimilated to different Polish vowel categories and which included height differences (results equal to or smaller than 3.37 on a seven-point Likert scale): /iː - ɪ/, /ɒ - ɑː/, /e - ʌ/, /ʊ - ɜː/, /e - ɪ/ and /e - æ/.

One could expect, on the basis of Bohn's (1995) desensitization hypothesis, that duration/tenseness differences would enhance the perception of vowels as different, however this did seem to be the case. It is noteworthy that the least similar vowel pairs, that is /e - æ/, /e - ɪ/ and /e - ʌ/ do not have duration/tenseness distinctions. (Dis-)similarity ratings seem to be based on assimilations to categories and height differences.

3.11. Conclusions

The four research questions about the perception of 11 British English monophthongs by Polish advanced learners of English in a formal instruction setting were answered on the basis of a longitudinal study.

As an answer to research question number one about the changes in assimilation patterns of English vowels to Polish vowels between stage one and stage two, we can conclude that for some vowels we observed progress toward new category formation, namely decreased goodness ratings for /æ, ʌ, ɑː/ and /ɪ/ (which is in line with Flege 1991). Further, the English /ɜː/ and /ɔː/ were considered dissimilar to their Polish counterparts in both stages, whereas the English /e/ and /ɒ/ were considered similar. When both spectral differences and duration differences were present between English and Polish vowels, they were assimilated with considerably lower goodness ratings. It seems that in the case of inherent duration differences between Polish and English vowels and greater acoustic difference between the vowels, the English vowels were more likely to be perceived as different from Polish vowels. This provides some support for the role of selective attention to features in L2 speech perception – subjects were able to observe and react to subphonemic differences between vowels.

With regard to research question number two, we can claim that identification results can well be related to assimilation results, with very detailed adjustments based on goodness rankings. Vowels assimilated as Single Category were identified with the lowest accuracy, whereas vowels from Cross Boundary assimilation types were identified most accurately. Depending on differences in goodness rankings, some Uncategorized vowels and vowels coming from Category Goodness assimilation types were identified at similar levels. Higher identification scores were obtained for vowels with larger differences in Category Goodness assimilation types. Moreover, vowels from Single Category and Category Goodness assimilations with relatively small differences in goodness ratings were better identified when the goodness ratings for a given vowel were lower.

Generally, long/free vowels were better identified than short/checked vowels, but the results were mediated by phonotactic restrictions and similarity to Polish vowels. Moreover, vowels from contrasts based on new tongue advancement distinctions were difficult to identify. Misidentifications rarely targeted a vowel with a different length. All these small adjustments to reliance on assimilation types in identification (i.e. reliance on categories) signal that listeners were attentive to phonetic details, both based on universals (longer duration is more conspicuous, tongue advancement distinctions are less frequent in world languages than tongue height distinctions) and selective attention to features (the larger the difference in Category Goodness assimilation, the easier the identification of the vowels coming from such a pair).

Research question number three asked about discrimination rates and potential changes. Since for numerous pairs discrimination was very good to excellent, there was little room for improvement here. Nevertheless, out of average and weakly discriminated pairs, significant improvement in discrimination was only recorded for the /uː - ʊ/ contrast, where both vowels were more consistently assimilated to the Polish /u/ in stage two.

With regard to research question number four, which asked about (dis-)similarity ratings and changes between stage one and stage two, as well as relationships with assimilation and discrimination results, it can be claimed that the most similar vowels were also poorly discriminated. The most dissimilar vowel contrasts came from Two Category assimilations and included the difference in the position of the tongue. Duration differences did not enhance judging vowels as dissimilar.

The universal tendency of vowel systems to have more tongue height than advancement distinctions is confirmed here: tongue height differences were more discernible than tongue advancement differences and vowels involving tongue height differences were perceived as more dissimilar than those which differ in tongue advancement. This result might have implications for teaching pronunciation. More attention should be paid to discriminating and practicing these vowel contrasts which rely primarily on tongue advancement: especially low vowels and high central vowels.

English low vowels were difficult for Polish subjects to perceive in all the tasks. They were categorized as the Polish /a/ in the assimilation task. The more retracted the stimulus was, the lower the goodness ratings were, especially in the case of /ɑː/ which is also a long/free vowel. In the identification task, the accuracy for low vowels spanned from very poor to medium. In the (dis-)similarity rating task, /æ - ʌ/ and /ɑː - ʌ/ were judged to be the most similar of all the pairs presented. Taken together, the results imply that /æ, ʌ/ and /ɑː/ were perceived by Polish learners in terms of the Polish /a/, and that the tongue advancement differences between them and longer duration of /æ/ and even more so of /ɑː/ were not enough to yield easy discrimination, identification or differentiation between these three vowels.

Contrasts involving duration/tenseness differences were neither easily discriminated nor prone to be rated as different. Duration/tenseness differences only contributed to lower goodness ratings of the free/long vowels. Whenever there were two English vowels categorized as one Polish vowel, it was the longer stimulus that received lowered goodness ratings. Three long vowels /iː, ɔː/ and /uː/ were well identified and /ɑː/ was better identified than other low vowels. Nevertheless, there were counter-examples: /ɜː/ was not particularly easy to identify (admittedly, it is a rare vowel with peculiar spectral characteristics) and /æ/ was the worst identified of all the vowels. The role of duration, however, mediated misidentifications: short vowels were mistaken for other short vowels and long vowels for other long vowels. The above results suggest that longer duration is a readily available cue of English vowels for Polish subjects in forced-choice identification, categorization and rating tasks, but it does not seem to be used as a cue in vowel discrimination or (dis-)similarity rating. These findings impose certain limits on the desensitization hypothesis (Bohn 1995) and supplement the results obtained by Bogacka (2004), who found that duration was more important than spectral differences in identification of /iː - ɪ/, Waniek-Klimczak (2005), who found that short and long vowels were differentiated by Polish learners of English, but to a lesser extent than by native speakers, and Rojczyk (2010b and 2011), who found that the /æ - ʌ/ contrast was primarily distinguished on the basis of vowel length. The results support the feature hypothesis (McAllister et al. 2002) and selective attention to features hypothesis (Pajak and Levy 2014), in that the Polish listeners with no experience of contrastive duration in the L1, find it difficult to perceive the cue in the L2. Moreover, they suggest that in some tasks the developing sensitivity to duration might be captured earlier than in other tasks.

We cannot claim that selective attention to features guides L2 speech perception development, but we observed that it supplements categories as the main elements in the system, especially when subjects assigned goodness ratings and found it easier to identify a given vowel thanks to a specific feature. Discrimination and (dis-)similarity rating results seem to have stemmed primarily from assimilation types. Familiar tongue advancement positions were accurately perceived when coupled with familiar length. In the case of tongue advancement distinction at unfamiliar heights, such as English low vowels and high central vowels, challenges in perception were easy to track.

Chapter Four

Perception of third and foreign language vowels

The next level of testing for the hypothesis on selective attention to features involves the domains of L3 and foreign[1] vowel perception by multilinguals. Foreign vowel perception means the perception of vowels of a language which the subjects had not been learning, whereas L3 vowels mean the vowels of a language the subjects had been learning as a second non-native language. Section 4.1. introduces theoretical attempts to account for multilingual influence in L3 acquisition, including the existing models of L3 acquisition, their verification in L3 speech studies by Wrembel (2015), and the Natural Growth Model (Dziubalska-Kołaczyk 2016, Dziubalska-Kołaczyk and Wrembel 2017). Section 4.2. presents previous research on non-native vowel perception with respect to three aspects. Section 4.2.1. shows that listeners with various L1s perceive non-native vowels differently. Section 4.2.2. analyzes the influence of inventory size on non-native vowel perception, whereas section 4.2.3. shows that native and non-native dialects play a role in non-native speech perception, in accordance with the role of acoustic detail in speech perception. Finally, section 4.2.4. discusses non-native perception by bilinguals. Section 4.3. begins with the presentation of the few existing studies on speech perception in L3. A summary of the assumptions of the present study is presented in section 4.3.2. and the aspects of selective attention to features which are relevant to the present study are described in section 4.3.3. The characteristics of vowel systems in the languages employed in the study are presented in section 4.4. The remaining part of the chapter is devoted to the study two (section 4.5.): research questions, method, results and discussion.

[1] In language teaching literature the term *foreign* might refer to the language studied in classroom. In speech perception literature usually the term *non-native* is used. Since all the languages discussed in this chapter, except for Polish, are non-native for the subjects, and we still need to distinguish between the languages the subjects are familiar with (L1, L2, L3) and the language they never had contact with, the term *foreign* is used to designate the latter.

4.1. Multilingual influence in L3 acquisition

In this section, aspects specific to L3 acquisition are presented to emphasize the fact that language acquisition by bilinguals differs from language acquisition by monolinguals. Section 4.1.1. discusses models of L3 acquisition while section 4.1.2. shows how these models were challenged by L3 studies conducted by Wrembel (2015), and as a result Dziubalska-Kołaczyk (2016) and Dziubalska-Kołaczyk and Wrembel (2017) proposed the Natural Growth Model (presented here in section 4.1.3.).

4.1.1. Models of L3 acquisition

Several models of L3 acquisition relating mostly to morpho-syntax have so far been proposed: the Cumulative Enhancement Model (Flynn, Foley and Vinnitskaya 2004), the L2 Status Factor Model (Bardel and Falk 2007 and 2012), the Typological Primacy Model (Rothman 2011 and 2015) and the Linguistic Proximity Model (Westergaard, Mitrofanova, Mykhaylyk and Rodina 2017).

The Cumulative Enhancement Model (Flynn, Foley, Vinnitskaya 2004) assumes that all languages known to a learner influence subsequent language acquisition if they can play a facilitative role, because all language learning should be accumulative and non-redundant. Neither the L1 nor the L2 maintains a privileged role. For the sake of cognitive economy, the parser should determine what is facilitative and what is not. Yet the mechanism that allows for facilitative transfer is unclear. Rothman and Halloran (2013: 57) observe that in the Cumulative Enhancement Model the facilitative transfer could only occur on "a property-by-property basis"[2], as only such a mechanism could guarantee that transfer from either the L1 or the L2 could be blocked if it were non-facilitative. Rothman and Halloran (2013) further criticize this idea on the grounds that it is too complicated and uneconomical from the cognitive point of view. To support their model, Flynn, Foley and Vinnitskaya (2004) presented data combing the production of three types of restrictive relative clauses in children and adults acquiring an L3.

---

[2] This would be in line with the hypothesis in the present study that specific features, whose use is contrastive in one of the languages known to a learner, can be re-employed in another language. In syntax, however, this idea is problematic, because it would imply that transfer in multilingualism is different from transfer in L2 acquisition, where full transfer at the initial state is widely accepted.

In the L2 Status Factor Model, Bardel and Falk (2007, 2012) and Falk and Bardel (2010, 2011) argue that syntactic structures, as opposed to vocabulary acquisition, are more easily transferred from the L2 rather than from the L1 in the initial state of L3 acquisition. The strong role of the L2 for both facilitative and non-facilitative transfer should be due to heightened metalinguistic knowledge that is common for L2 and L3 acquisition in the formal learning setting. In the L2 Status Factor Model, L1 transfer is allowed only when the learners have high explicit metalinguistic knowledge of the L1.

According to Rothman's (2010, 2011, 2013 and 2015) Typological Primacy Model, structural similarity determines whether the L1 or the L2 is the source of transfer. Neither the order of acquisition nor facilitativeness plays a role. Relying on general linguistic economy and cognitive-processing factors, the Typological Primacy Model argues that one of the two previously acquired languages will be transferred completely in the initial stages (Rothman 2013). The internal parser should decide which language resembles the L3 more on the basis of the underlying linguistic similarities between: the lexicon, syntactic structure, functional morphology and phonological or phonotactic cues. The parser filters the L3 input, making structural comparisons in the above listed order and chooses to completely transfer either the L1 or the L2. As in the case of the Cumulative Enhancement Model, either the L1 or the L2 might influence the L3, but the Typological Primacy Model predicts that transfer will then be both facilitative and non-facilitative. This is the consequence of choosing one of the previously learned languages in its entirety as a source of transfer in the initial stages of L3 acquisition.

In the Linguistic Proximity Model, Westergaard et al. (2017) propose that all previously acquired languages remain active throughout the learning process and that third language acquisition (actually it is L_n) is based on gradual property-by-property learning. Instead of a holistic transfer based on typological similarity between the L3 and one of the previously acquired languages, Westergaard et al.'s (2017) model assumes cross-linguistic influence when a property present in the L3 is structurally similar to a property from one of the languages the learner already knows. Under this assumption, cross-linguistic transfer of a given property happens when the learner has the ability to parse the input and deconstruct the abstract linguistic properties. Interestingly, the Linguistic Proximity Model rules out any influence of other languages at the early stage of L3 acquisition, as there has been no empirical evidence of complete transfer. Westergaard et al. (2017) criticize Rothman (2015) for his proposal that because complete transfer is motivated by cognitive economy, its aim is to avoid redundancy in learning and therefore transfer should happen as early and as completely as possible to fulfill its role. They argue that incremental learning is more economical because one only transfers those properties which one knows are relevant. This strategy spares effort which would be needed to unlearn all the properties which turn out to be irrelevant. In addition to facilitative influence based on structural similarity or overlap between grammars, non-facilitative transfer is also predicted by the model. Non-facilitative transfer takes place when a learner mistakenly interprets the L3 input and assumes that a property is similar to a property from a previously acquired language.

The three models were developed on the basis of morphosyntactic studies. Since they generate conflicting predictions, Wrembel (2015) verified these three models when examining L3 speech.

4.1.2. Complex nature of the cross-linguistic influence in L3 speech

Testing the proposed models of multilingualism, Wrembel (2015) conducted three studies concerning L3 speech: accentedness ratings, VOT measurements and metaphonological awareness.

In study 1, raters were asked to evaluate the accentedness of L3 speech samples and identify the subjects' first language. The results revealed the prevailing effect of L1 and the weak, though consistent, effect of L2. The role of typological proximity was rather facilitatory than determining. The results also partially refuted Hammaberg and Hammaberg's (1993 and 2005) hypothesis that L2 plays a more dominant role than L1 in L3 acquisition. The study confirmed De Angelis' (2007) assumption about the combined cross-linguistic influence.

Wrembel's (2015) second study consisted of measuring VOT values in L1, L2 and L3 speech. The participants differentiated their VOT values depending on whether they spoke L1, L2 or L3. The new 'hybrid' L3 categories deviated from both the L1 and L2 categories, but at the same time they were compromise values, transferred strongly from the L1 Polish and more moderately from the L2. Again, a combined cross-linguistic influence (De Angelis 2007) was confirmed. Both facilitative and non-facilitative transfer occurred. As opposed to the VOT hybrids between the native and target values reported in the L2 literature (Flege 1987b; Flege and Eefting 1988), the L3 VOT hybrid seemed to be influenced by native VOT values, target VOT values and the L2 VOT specifications. Some facilitating effect of typology was also reported.

The third study in Wrembel (2015) examined metaphonological awareness elicited through verbal protocols. The subjects declared transfer from both L1 and L2 to L3, although the majority indicated that L2 transfer prevailed, lending support to the foreign language effect or the L2 status (Cenoz, 2001; Hammaberg and Hammaberg, 2005; Bardel and Falk 2007) or a combined cross-linguistic influence (De Angelis 2007). Also, comments on multilingual advantage in L3 learning were attested.

Regarding potential sources of multilingual transfer, Rothman (2015: 182) identified four logically possible scenarios: (1) no transfer, (2) absolute L1 transfer, (3) absolute L2 transfer, and (4) L1 and/or L2 transfer. Wrembel's (2015) results clearly point to both L1 and L2 as sources of transfer, although estimating the proportions of the two is in no way possible at this point. The Cumulative Enhancement Model for Language Acquisition (Flynn et al. 2004) predicts that the transfer of properties from L1 and L2 can only manifest itself when it is facilitative for L1 acquisition. The model also predicts that non-facilitative transfer should be blocked. The results presented by Wrembel (2015) include examples of both facilitative and non-facilitative transfer. The L2 Status Factor Model (Bardel and Falk, 2007, 2012 and Falk and Bardel, 2011) assumes that the L2 is the prevailing source of transfer. The assumption about the greater cognitive similarity between L2 and L3 was indeed evident in the metalanguage awareness study, but globally the effects of L2 did not prevail over the effects of L1. Rothman's (2010, 2011, 2013, 2015) Typological Primacy Model assumes access to both L1 and L2 in the initial state, but that transfer occurs holistically from either L1 or L2, depending on the perceived typological proximity of the languages involved. This claim is motivated by the principle of a general cognitive economy demanding that the least effort is dedicated to a given cognitive task. Wrembel's (2015) data did not support the hypothesis of full transfer from one language only. Many cases of a property-by-property or gradual transfer from both languages were evident[3]. As Wrembel (2015: 405) observes, it is unrealistic to expect that inhibition processes should completely suppress the activation of the typologically more distant language. Some evidence that her results provide in favor of typologically-based transfer is inconsistent. Unlike the Cumulative Enhancement Model, the Typological Primacy Model allows for both facilitative and non-facilitative transfer, which was so often present in Wrembel's (2015) data. Disentangling the factors of typology and L2 status turned out to be to impossible in the case of Wrembel's (2015) studies. Only some of the results could be attributed to typological facilitation.

4.1.3. Natural Growth Model: a reminder about the crucial questions

In second language perception studies, the focus has so far been on the L1 filter, means of overcoming this and reaching native-likeness in perceiving second language contrasts. Accordingly, studies have been bound to focus on phonetic details such as VOT and only those differences in formant values or duration etc. which are noticeable. So as not reduce the results of second/third language acquisition studies to "a collection of observable or elicited details", it is worth noting a recent proposal to view second and third language acquisition in terms of the Natural Growth Model (Dziubalska-Kołaczyk 2016, Dziubalska-Kołaczyk and Wrembel 2017). The model has the ambition to explain all the crucial aspects of speech acquisition, including L1, L2, L3, cross-linguistic influence, language attrition and death, and at the same time be interdisciplinary. The proposed multilingual acquisition scenario encompasses the following postulates: 1) to study phonological representations or processes per se in addition to phonetic categories or bundles of features 2) to apply the concept of universals as typological and/or statistical preferences which could explain the reported asymmetries in perception. The Natural Growth Model of Acquisition (Dziubalska-Kołaczyk 2016, Dziubalska-Kołaczyk and Wrembel 2017), which is based on Natural Complexity Theory (Dziubalska-Kołaczyk in preparation), assumes the gradual emergence of an L3/Ln phonology influenced by L1, L2, other languages, typology, universals and context. The estimated impact depends on the frequency and type of language usage. Both context

[3] This could now be interpreted as evidence in favor of the Linguistic Proximity Model (Westergaard et al. 2017).

internal and external processes are claimed to be 'vehicles' of the growth. The postulated gradual emergence is dependent on multiple factors and, as such, it is a step forward in comparison to previous L3 morphosyntactic models, which assume the precedence of either L1 or L2 over the other.

In fact, however, the Natural Growth Model of Acquisition (Dziubalska-Kołaczyk 2016, Dziubalska-Kołaczyk and Wrembel 2017) as it was presented, is not yet a model which generates testable predictions, in the sense the Perceptual Assimilation Model and the Speech Learning Model do. At present, its main strength lies in offering post-factum explanation. Nevetheless, Dziubalska-Kołaczyk and Wrembel (2017) challenge linguists working on speech by asking whether we are still looking for the bigger picture, an explanatory framework when interpreting empirical data, and whether we are seeking an explanation rather than concentrating on small details. They provide an incentive to rethink the goals of current phonetic research in second and third language acquisition. Certainly, we may find numerous studies that simply concentrate on small phonetic details or test the same models against different language combinations, experimental conditions or training methods, but they do not contribute to falsification or elaboration of the models. They multiply experimental designs, but they do not advance the theory. Yet, in order to advance the model or theory, one crucially needs detailed, targeted studies which address the loopholes in the theory. Usually, only a very meticulous study advances the model just one step forward, and leaves many questions unanswered, or better still inspires new questions. Asking the big questions and answering them with a single study is not feasible. For example, as a reaction to Port and Leary's (2005) paper, how could one prove or falsify a phoneme? Concluding, new models, or big proposals such as the Natural Growth Model, are most welcome, but detailed work is likewise needed to verify them. The present studies are an answer to the call for studies of multilingual influence in various configurations (i.e. with different groups of learners, order of acquisition, acquisition settings, and levels of proficiency) to advance the complex theory of third language acquisition.

4.2. Cross-language perception of non-native vowels

Previous studies suggest that non-native vowel perception depends on the listeners' L1 or dialect. There is also fragmentary evidence that it might vary as a function of any previously acquired languages. The following subsections will analyze which factors influence non-native vowel perception. Subsection 4.1.1. presents non-native vowel perception by listeners with various L1s to show that the native inventory crucially determines non-native speech perception. The next subsection examines the relationship between the inventory size in L1 and non-native vowel perception. To strengthen the argument about the determining role of subphonemic acoustic details in speech perception, subsection 4.1.3. presents the role of native dialect.

4.2.1. Non-native vowel perception by listeners with various L1s

It is well attested that a listener's L1 influences L2 vowel perception. For example, Bohn and Flege (1990) showed that German listeners will not find distinguishing between English /iː/ and /ɪ/ challenging, presumably because German has a similar contrast, whereas the English contrast /e/-/æ/ is more challenging, presumably because /æ/ is not a part of the English inventory. This section reports on two studies which examined how users of various L1s perceive non-native sounds differently.

Best, Halle, Bohn and Faber (2003) examined cross-language perception of non-native vowels to investigate the phonological and phonetic effects of listeners' native languages. They tested the perception of Norwegian vowels by English, French and Danish listeners. All the four languages use /i/ and /u/, all except English have /y/, but only Norwegian uses /ʉ/. Moreover, the vowels /i, y/ and /u/ are realized differently in the languages in question. In the experiment, American, Danish and French listeners discriminated Norwegian /i - y/, /y - u/, /y - ʉ/ and /ʉ - u/ vowel pairs and categorized the vowels involved. The results coincided with the languages' phonological and phonetic properties.

The results of the categorical AXB discrimination test were analyzed using a language x contrast ANOVA. There was a significant language effect, which showed that Danish listeners (98% correct discrimination) were more accurate than French listeners (96%) and then American listeners (92%). The contrast effect indicated that discrimination of /iː - ɪ/ (88.5% correct) was worse than discrimination rates for the other three contrasts, which ranged from 97 to 99% correct. Finally, a significant language x contrast interaction showed that Danish listeners discriminated all the contrasts near to the ceiling (97-98% correct). So did French and American listeners in the case of /y - u/, /y - ʉ/, /ʉ - u/ contrasts (98-99% and 96-98% respectively). The /i - y/

contrast was discriminated significantly less well by the French listeners (89%), and significantly worse by the American listeners (79%).

In the categorization task, Norwegian /y - u/ were assimilated in terms of Perceptual Assimilation Model's (Best 1995) Two Category native contrast. Danish and French listeners also assimilated /ʉ - u/as a Cross Boundary contrast. The /i - y/ was assimilated as a Two Category contrast by Danish listeners, as Category Goodness by French subjects and as a Single Category by American English listeners. Both the French and American English listeners assimilated /y - ʉ/ as a Two Category native contrast, but Danish listeners assimilated it as a Single Category contrast. The authors, Best et al. (2003), concluded that the results were largely in line with the phonologically contrastive and non-contrastive phonetic-articulatory properties of the listeners' L1s.

Iverson and Evans (2007) examined whether L2 English users with different L1s (Spanish, French, German and Norwegian) use different cues (e.g. formant movement or vowel duration) in English vowel perception. The study used an array of tasks: natural English identification; identification of English vowels with flattened formant movement and equated duration; perceptual mapping of L1 and L2 synthetic vowels in a five-dimensional vowel space and natural English vowel assimilation to L1 categories. Although the subjects with richer vowel inventories (German and Norwegian) recognized English vowels more accurately than subjects with smaller L1 vocalic systems (French and Spanish), they did not differ significantly in how they perceived English vowels. All groups of subjects used both formant movement and duration in English vowel perception and showed signs of learning aspects of a new vocalic inventory rather than just assimilating the English sounds to native categories. Even in the case of L1 German or Norwegian, when assimilation to L1 categories would be sufficient to distinguish between L2 categories, the subjects learned secondary cues (formant movement and duration) instead of simply adding new primary cue specifications for static F1/F2 targets.

These two studies by Best et al. (2003) and Iverson and Evans (2007) prove that it is both the comparison between the L1 and L2/FL as well as the relationships between the non-contrastive phonetic properties of sounds that determine non-native speech perception.

4.2.2. Non-native vowel perception and L1 inventory size

Some previous research suggested that the size of L1 and L2 vowel inventories might predict non-native and L2 vowel perception (Fox, Flege and Munro 1995, Lengeris 2009, Bundgaard-Nielsen et al. 2011). The idea was that listeners with smaller L1 vowel inventories find perceiving vowels from larger L2 inventories difficult, because more than one L2 vowel needs to be mapped onto some L1 categories. By the same token, a larger L1 inventory should facilitate perception of vowels from a smaller L2 inventory. Contrary to this view, the Second Language Linguistic Perception Model (Escudero 2005, 2006, 2009) states that non-native and L2 vowel perception is predicted by L1-L2 acoustic relationships. Elvin, Escudero and Vasiliev (2014) tested whether foreign vowel discrimination accuracy is better predicted by vowel inventory size or cross-linguistic acoustic properties, pointing to the latter. Naïve Australian English and Iberian Spanish listeners were presented with six Brazilian Portuguese vowel contrasts. The Australian English vowel system includes all the Brazilian Portuguese contrasts, whereas Iberian Spanish has only five vowels (it lacks the mid vowels present in Brazilian Portuguese), which are, however, acoustically more similar to Brazilian Portuguese vowels. The relative weight of vowel inventory size and acoustic proximity of the native and non-native vowels still needed to be determined.

Accordingly, Alispahic, Mulak and Escudero (2017) examined the effects of vowel inventory size and acoustic properties on non-native vowel perception. They tested discrimination and categorization of five Dutch vowel contrasts by Australian English listeners, whose L1 has more vowels than Dutch and Peruvian Spanish listeners, whose L1 has fewer vowels than Dutch. They found no effect of L1 inventory size. Participants in both groups were better at discriminating contrasts which were predicted as being easy based on L1-L2 acoustic relationships, and worse at discriminating contrasts which were predicted as being difficult based on L1-L2 acoustic relationships. The results also showed that subjects with a larger L1 inventory activated numerous native categories for some Dutch vowels, whereas listeners with a smaller inventory activated fewer categories and consequently had higher accuracy scores. Alispahic et al. (2017) concluded that non-native vowel perception is better predicted by L1-L2 acoustic relationships than by inventory size alone.

Although the inventory size might matter in non-native speech perception, acoustic proximity between the L1 and non-native vowels is also a crucial factor in non-native vowel perception.

4.2.3. Non-native vowel perception and native dialect

The studies reported above showed that acoustic proximity between the L1 and non-native sounds is crucial for determining non-native speech perception. Several recent studies have also shown that the perception of non-native sounds can be affected by the native language dialect.

The Second Language Linguistic Perception Model (Escudero and Boersma 2004, Escudero 2005) assumes that listeners' perceptions of native and non-native sounds match the acoustic properties of relevant sounds in their native dialects. Escudero and Boersma (2004) showed that native speakers of English perceive the same /i/ and /ɪ/ tokens differentially, depending on whether they are native Scottish English or native Southern British English listeners. Although they did not control for native dialect, Mayr and Escudero (2010) reported that for L2 listeners variation in responses to German front rounded vowels could potentially be attributed to their Southern or Northern British English dialects. Chládková and Podlipský (2011) examined how Bohemian Czech and Moravian Czech listeners perceptually assimilated Dutch vowels. These two dialects realize the /iː - ɪ/ contrast differently. In Bohemian Czech /iː/ has a lower F1 and is longer than /ɪ/, whereas in Moravian Czech it is primarily the length difference that distinguishes the two vowels. Consequently, the Dutch /i/ is assimilated to /iː/ by Bohemian Czech listeners and to /ɪ/ by Moravian Czech listeners. Another example of a study in which the L1 dialect affected L2 perception was a study by Escudero, Simon and Mitterer (2012). Northern Dutch and Flemish speakers, who realize the Dutch vowels /ɪ, ɛ, a/ and /ɑ/ differently, also perceived the Southern British English /e/ and /æ/ differently. Similarly, Escudero and Williams (2012) found that acoustic differences between Peruvian Spanish and Iberian Spanish vowels result in differential L2 Dutch vowel perception. Also, the acoustic properties of the target dialect have been shown to determine L2 speech perception: Spanish learners classify English /i/ and /ɪ/ as Spanish /i/ and /e/ if they learn Scottish English and as Spanish /i/ if they learn Southern British English. In conclusion, it can be assumed that it is the relationship between the acoustic properties of both

L1 and L2 vowels which determines L2 vowel assimilation patterns to L1 categories.

4.2.4. Non-native perception by bilinguals

Research on non-native vowel perception by bilinguals aims at answering questions about the advantage of bilinguals over monolinguals or the source and nature of the bilingual advantage.

An early study by Cohen, Tucker and Lambert (1967) examined the repetition of phoneme sequences by English and French monolinguals, bilinguals dominant in either language and early bilinguals. All groups were most accurate when asked to repeat sequences occurring in both languages, less accurate with sequences occurring in just one of them and the least accurate with sequences occurring in neither language. Bilinguals outperformed monolinguals. Attested distortions happened in the direction of the L1.

Rochet (1995) studied how Brazilian Portuguese, Canadian English and French listeners labelled Canadian English vowels. Brazilian Portuguese and Canadian English participants were asked to identify synthetic vowels on a high vowel continuum as /i/ or /u/ and French listeners as /i, y/ or /u/. The vowels the French listeners labeled as /y/ were mostly labeled as /u/ by Canadian English listeners and as /i/ by Brazilian Portuguese listeners. The study demonstrates that language background influences not only the goodness rating of a vowel as similar to the L1, but also which native language categories a given non-native vowel is assigned to.

Knowing that monolingual listeners depend on their L1 when categorizing and discriminating unfamiliar non-native contrasts, Antoniou, Best and Tyler (2013) investigated whether early bilinguals are constrained in the same fashion by their languages or whether they possess a cumulative bilingual advantage. They examined the perception of Ma'di stops by Greek and English bilinguals and monolinguals. Ma'di stop voicing distinctions exist in neither Greek nor English. In line with the predictions, English monolinguals assimilated Ma'di prevoiced plosive and implosive stops and the coronal voiceless stop to English voiced stops. The Greek monolinguals assimilated the Ma'di short-lag voiceless stops to Greek voiceless stops and the prevoiced implosives and coronal stops as Greek voiced stops. Consequently, the English monolinguals poorly discriminated the non-native voicing contrasts, while the Greek monolinguals did it very well. To manipulate the language mode, bilinguals were instructed either in English or in Greek. Depending on the language mode, they categorized Ma'di stops similarly to respective monolinguals. Their discrimination, however, did not depend on language mode. They were intermediate to the two monolingual groups for the prevoiced-voiceless contrast. Bilinguals' and monolinguals' discrimination of prevoiced plosive-implosive contrasts was equally poor. The authors concluded that there was no bilingual advantage for unfamiliar non-native contrasts. The results nevertheless proved that bilinguals were uniquely configured language users, who differed from either monolingual group. It was also evident that categorization was more influenced by bilingual processing than discrimination.

Concluding, foreign speech perception research on bilinguals has been fairly limited. No conclusive results as to the bilingual advantage for non-native contrasts have been obtained. There still exists the need to answer questions about the source and limits of transfer from the L1 and L2 to a non-native language in speech perception.

4.3. Speech perception in L3: questions about broad-based bilingual advantage vs. a narrow L1/L2 to L3 transfer

This section presents previous research in L3 speech perception, assumptions related to multilingualism in the present study and the selective attention to features hypothesis for the L3 and foreign language perception.

4.3.1. Previous research on perception in L3

Perception in the L3 has not been widely studied so far with the exception of studies by Werker (1986), Patihis, Oh and Mogilner (2015), Kopečková (2015), Cabrelli Amaro (2016), Onishi (2016), and Wrembel, Marecka and Kopečková (under review). These studies offered no conclusive results as to the direction of transfer or the role of features. One of the important issues in the perception of unfamiliar sounds by bi- and multilinguals is whether they are better than monolinguals. If so, the question remains whether the advantage is broadly based on more flexible cognitive facility or whether bilinguals simply benefit from exposure to L2 sounds with similar distinctions to the ones tested in an L3 or an unfamiliar language in a particular study.

Werker (1986) designed a study to test whether discrimination depends on specific linguistic/perceptual experience, or whether extended but non-relevant linguistic experience can promote broad perceptual flexibility. Bilingual and trilingual subjects were compared to monolingual English listeners on their ability to discriminate phonetic contrasts that were not present in any of their native language(s). The results suggested, however, that extended, nonspecific linguistic experience did not contribute to broad perceptual flexibility.

For Patihis, Oh and Mogilner (2015), broad-based bilingual advantage could also be due to enhanced flexibility in phoneme categorization. They designed a study to determine whether multilingual discrimination of foreign contrasts is characterized by a broad-based bilingual advantage or a narrow L1/L2 to L3 transfer. English monolinguals, English-Spanish and English-Armenian bilinguals and trilinguals, who had never been exposed to Korean, were asked to discriminate Korean stops. Spanish English bilinguals performed just as well as English monolinguals and these two groups were outperformed by Armenian-English bilinguals. Patihis et al. (2015) found no broad bilingual or trilingual advantage in discriminating the sounds of an unfamiliar language. The advantage in perception was limited to the bilinguals who knew Armenian, which has similar phonological distinctions to the Korean ones tested in the study. Experience with a distinction between aspirated and non-aspirated stops in Armenian resulted in enhanced perception of Korean aspirated and non-aspirated stops. The authors concluded the paper by claiming that the bilingual advantage they found was not due to greater flexibility in categorizing phonemes, but due to narrow L1/L2 to L3 transfer, facilitated by experience with phonemic contrasts present simultaneously in the target language and one of the languages well known by the subjects.

A study by Antoniou, Liang and Ettlinger and Wong (2015) was a phonetic, but not a perception study. Nevertheless, it has important implications for bilingual advantage and the role of features. The study examined the interaction between the factors which are supposed to be advantageous for learning a non-native language: being bilingual, learning a related language and having to learn unmarked, rather than marked, features. The subjects were asked to learn artificial language words which were differentiated on the basis of phonetic contrasts. Bilinguals outperformed monolinguals in the case of unmarked features, but for difficult contrasts only, the subjects with language specific experience performed better.

Kopečková (2015) found that children with L1 Polish and L2 English were more sensitive to the differences between Polish and English vowels if they also learned German, French or Irish as L3 as opposed to a situation where they were only exposed to Polish and English. Since the examined sound contrasts exist both in the subjects' L2 and L3s, that study could not conclusively determine whether the enhanced sensitivity could be ascribed to the cognitive flexibility of L3 learners or simply a narrow positive transfer between L2 and L3.

Cabrelli Amaro (2016) tested the Phonological Permeability Hypothesis (Cabrelli Amaro and Rothman 2010), which predicts that adult L2 phonological systems are less stable than L1 systems. To this end, L3 Brazilian Portuguese regressive transfer in perception by L1 English/L2 Spanish and L1 Spanish/L2 English listeners was examined. A forced-choice preference task was used to test the subjects' preference for Brazilian Portuguese or Spanish vowel allophone in a word-final position (in Brazilian Portuguese, but not in Spanish, vowels are reduced in a word-final position). Neither experimental group showed any vulnerability to L3 Brazilian Portuguese influence. The subjects preferred Spanish-like fully realized word-final vowels and they made their decisions equally fast. Summarizing, there was no difference between L1 and L2 Spanish perception data, though parallel production data for L2 Spanish deviated from L1 Spanish and Spanish monolingual controls, providing preliminary support for the Phonological Permeability Hypothesis.

Onishi (2016) examined the influence of experience in the L2 on perception in the L3 at the beginner level. The author tested Korean native listeners' perception of L2 English and L3 Japanese sounds and argued that the better the L2 perception, the better the L3 perception. One methodological issue, which might partially undermine the conclusions, is that two different, rather small groups of Korean listeners were employed in the identification task and in the discrimination task. As a result, we cannot compare the results of the two tasks directly, as any differences may stem from the fact that different participants took part in the tasks. Worthy of note is that a categorical AXB discrimination task, which was used as a discrimination task, is actually very demanding on the listeners as they need to focus on relevant phonological features only and disregard inter-speaker variation when three different speakers are used in a triad.

Wrembel, Marecka and Kopečková (under review) found that L3/Ln Polish vowels were perceived by adolescent learners as slightly more similar to L2 English rather than to L1 German categories. The difference was not large, but it was taken to suggest the L2 effect. In the same paper, sibilants were rather accurately discriminated (83.13% correct). The authors argue that high discrimination accuracy should be ascribed to the facilitation effect of multilingualism, or, in other words, broadly-based multilingual advantage. This claim is, however, questionable because no data for a control group of monolingual German listeners (certainly highly unlikely in this age group) were available. Also, the task employed, the AX discrimination task, is a rather easy discrimination task, which encourages reliance on purely acoustic information rather than phonological processing. Maybe the task was not challenging enough to promote larger differences between the groups. On the other hand, the discrimination rate of 83.13% is certainly above the chance level, but is also far from excellent. The groups were rather small – there were 10 participants altogether. The conclusion of this study about the facilitation effect of multilingualism certainly needs to be cross-checked in further research.

There have been few studies concerned with L3 perception, certain methodological shortcomings have appeared, and results related to the direction of transfer, the role of features and bilingual advantage are not exhaustive, so the field needs more controlled studies on L3 perception.

4.3.2. Assumptions for the current study

The present experiments will test whether listeners with a common L1 (Polish) are affected by the acquisition of L2 (English) or L3 (French or Dutch) when perceiving Dutch vowels (Dutch vowels were foreign, unfamiliar vowels for learners of English and French and L3 vowels for learners of Dutch).

The subjects in the present experiment knew at least two languages and were tested on foreign sound perception (from a language they had not had experience with) in the case of learners of L2 English and L3 French and were tested on L3 perception in the case of learners of L3 Dutch. All subjects knew English to some extent – the details are given in section 4.5.2.1. Hammarberg (2010) distinguishes between the L1 which is acquired before puberty, the L2 which is acquired after puberty and which allows the learner to develop strategies to learn foreign languages, and the L3 which is acquired when a person already has knowledge of one or more L2s. This terminology is considered superior to a linear model in which languages are numbered in the order of acquisition: the language acquired first is called the L1, the second acquired language is called the L2, and the third one – the L3. In the linear model the L3 could therefore refer to an L3 which is still a native language (as in the case of children brought up in bilingual families living in a community where yet another language is spoken), L3 as the first non-native language (for example, in the case of bilinguals who start learning another language at school) and L3 as the second non-native language (the usual understanding of the term in monolingual societies). As Hammarberg (2010) notices, the linear model misses important information about languages acquired simultaneously, with low proficiency, with limited types of knowledge, intermittently, or almost unintentionally due to language closeness. According to a different tradition, going back to the Critical Period Hypothesis (Lenneberg 1967), languages are considered to be L1s if they were acquired before puberty and L2s if they were acquired after puberty, so the order of acquisition of multiple L1s or L2s does not really matter (Cawalho and da Silva 2006, Davidiak 2010).

Although not falling directly into the domain of L3 acquisition, these experiments actually examine the interaction of multiple languages in speech perception and benefit from being grounded in L3 and multilingual literature. Traditionally, L3 is considered to be a language which a learner has already been learning and is familiar with at least to a certain extent (Hammarberg and Hammarberg 2005). The present study examines the influence of the sound system acquired during second and third language acquisition on the perception of L3 or foreign vowels (depending on the group of subjects). The tested vowels were Dutch vowels. Learners of Dutch obviously studied Dutch as the L3, and they were advanced learners, so we tested L3 perception in their case. For learners of English and French, the tested Dutch vowels were the sounds which they never had contact with. In the case of learners of English and French, this stage resembles what is called the *initial stage* in L3 acquisition (Rothman 2011), although this term presupposes that subjects intentionally begin to learn L3, which was not the case here.

Certain terms and assumptions stemming from research on L3 or multilingualism can inform the experiment proposed here. Firstly, the assumption of combined cross-linguistic influence (De Angelis 2007:21) is adopted. This term is based on the term cross-linguistic influence (Sharewood-Smith and

Kellerman 1986), which encompassed influences on the target language such as: transfer, interference, avoidance, borrowing and attrition. To acknowledge the possibility that in multilingualism the source of influence on the target language might simultaneously come from more than one language, De Angelis (2007) proposed the term 'combined cross-linguistic influence'. Some of the identified factors which determine multilinguals' reliance on previously learned languages include: language distance, target and source language proficiency, recency of use, length of residence, level of exposure, order of language acquisition and formality of situation.

Lastly, it can either be assumed after Ringbom (1987) and De Angelis (2007) that non-native languages can be sources of cross-linguistic influence irrespective of the level of proficiency, or it can be assumed after Gut (2010) that the proficiency threshold level in a non-native language must be sufficiently high in order to exert influence on another non-native language.

4.3.3. Selective attention to features: hypothesis for L3 and foreign speech perception

Bohn and Best (2012) and Pajak and Levy (2014) tested how features used in the L1 influence perception of non-native sounds. Bohn and Best (2012) specifically discussed the role of contrastive lip rounding. They concluded that better-than-native discrimination of English /w - j/ contrast by Danish and German listeners was due to the fact that both Danish and German use lip rounding contrastively to distinguish between high front rounded and unrounded vowels. Their hypothesis that the native vowel system influences non-native approximant perception prompted them to generate testable predictions. They hypothesized that listeners whose L1 uses contrastive lip rounding for vowels should discriminate American English /w - j/ comparably to French, Danish and German listeners. This scenario should be valid for Swedish and Norwegian listeners (/i, y, ʉ, u/ in the inventory) and Turkish listeners (/i, y, ɯ, u/ in the inventory), whose native languages, like Danish and German, have /j/ but not /w/. Bohn and Best (2012) also suggested that native speakers of languages with a rounding contrast only for back vowels, such as Korean or Portuguese (both have /i, ɯ, u/, but not /y/), be examined. For speakers of languages with non-contrastive lip rounding, but with counterparts to which English /w - j/ could be assimilated, such as Farsi, Hebrew and Czech with /i - u/ and /v - j/, they predicted that levels of discrimination would be comparable to those for English listeners, or potentially exhibit phonetic-level effects at the /w/ end of the continuum.

This study tests the hypothesis that selective attention to features applies in L3 and foreign language speech perception. To this end we will examine the extent to which selective attention to features is evident in L3 and foreign speech perception. We will check whether all languages known to the subjects, as opposed to the L1 only, influence the ways in which listeners discriminate and categorize L3 or foreign sounds.

4.4. Language characteristics: Vowel inventories of Dutch, English and French.

In order to examine whether non-native vowel perception depends on the specific L2 or L3 the subjects have been exposed to, and whether the selective attention to features can also be based on the L2 or L3, rather than only on the L1 as has been previously attested, the current study examines the perception of non-native Dutch vowels by L1 Polish learners of L2 English, L3 French or L3 Dutch. The choice of languages serves the purpose of examining the role of features related to acoustic characteristics of vowel quality, duration and lip rounding.

Dutch has an extensive vowel inventory with front unrounded, front rounded and back rounded vowels. Dutch vowels are also distinguished by length and tenseness. Adank, van Hout and Smits (2004) showed that the nine Dutch monophthongs /i, ɪ, ɛ, a, ɑ, ɔ, u, y/ and /ʏ/ can be relatively well separated on the basis of their steady-state characteristics, whereas both the long mid vowels /e, ø, o/ and the three diphthongs /ɛɪ, ɔu, œy/ crucially need to be supplemented with information on their dynamic characteristics. When it comes to length differences, Dutch vowels can be labelled as long vowels /a, e, ø, o, ɛɪ, ɔu/ and /œy/ and short vowels /i, ɪ, ɛ, ɑ, ɔ, u, y/ and /ʏ/.

British English has 11 monophthongs: /iː, ɪ, e, æ, ʌ, ɑː, ɔː, ʊ, uː, ɜː, ə/ (Wells 1962, 1982, 1984, Gimson 1964, 1980, 1984, Deterding 1990, Upton, Kretzschmar and Konopka 2003, Hawkins and Midgley 2005). It does not have front rounded vowels, but /uː/ and /ʊ/ are centralized and have little lip rounding. Cruttenden (2001) describes the fronting of /uː/ to /ʉ/ as well-established and the unrounding of /uː/ and /ʊ/ as a recent innovation. Hawkins and Midgley's (2005) acoustic study confirmed higher F2 for /uː/

and /ʊ/ and higher F1 for /e/ and /æ/. Duration or tenseness are considered to be secondary cues to English vowels.

French also has a relatively large vowel inventory: /i, y, e, ø, ɛ, œ, a, ɔ, o, u, ə/. It uses four height distinctions and three backness distinctions (Fougeron and Smith 1993). All its back vowels are rounded, whereas the front high, high mid and low mid pairs of vowels consist of unrounded and rounded vowels, but the rounded vowels are lower and centralized compared to the unrounded variants (Raphael, Bell-Berti, Collier and Baer 1979) The two central vowels are a mid schwa with some rounding and a low /a/. Contemporary French does not use duration contrastively.

It is considered to be vital that all the subjects are Polish learners of various second or third languages, rather than native speakers of respective languages, because a common L1 will allow for tertium comparationis. All subjects will identify the Dutch vowels in terms of Polish vowels. Repetitive as the design of the experiments might seem, its primary aim is to allow for comparison of vowel perception in terms of features and categories among various groups of learners.

The denser a vowel system in a given part of the vowel space is (the more phonetic features which are employed to express contrast in a given area), the more likely it is that other vowel contrasts will be perceived more accurately in this part of the vowel space, because more feature clusters should be at the listener's disposal. Learners of English were familiarized with English vowels – twice as many categories as Polish has. Learners of French should be better at perceiving front rounded vowel distinctions because French uses such vowels. Learners of Dutch should also cope well with front rounded vowels, and moreover with tongue advancement distinctions, as Dutch has more of these than either English or French. More detailed analysis will be offered on the basis of individual vowel features.

4.5. Study two

This section presents the second study in this book. The study examines L3 and foreign language vowel perception.

4.5.1. Research questions

In the present experiment, perception of Dutch vowels by Polish learners of English, French and Dutch was tested. Examining listeners' categorical discrimination and categorization of Dutch vowels permitted the two following major questions to be addressed and subsequent predictions to be made:

1. Does the listeners' L2 or L3 influence their perception of Dutch/foreign language vowels?

It has been well attested in the literature that non-native vowel perception depends on the listeners' L1s (Best et al. 2003), inventory size (Fox, Flege and Munro 1995, Lengeris 2009, Bundgaard-Nielsen et al. 2011), acoustic proximity of native and non-native vowels (Elvin et al. 2014 and Alispahic et al. 2017), dialect (Chládková and Podlipský 2011, Escudero et al. 2012, Escudero and Williams 2012), and bilingualism (Antoniou et al. 2013). Antoniou et al. (2013) focused on cumulative bilingual advantage in the perception of stop voicing distinctions. They found that categorization depended on language mode, but that discrimination did not. No bilingual advantage was found for unfamiliar contrasts. Generally, categorization was influenced more by bilingual processing than discrimination was. This study checks whether Polish listeners with various L2s (English, French and Dutch) perceive Dutch vowels differently. In the light of previous research on L3s in other linguistic domains, the L2 could directly influence foreign vowel perception. Also, in the L2 speech perception literature, vowel perception seems to depend on the listeners' experience with L1 vowel inventory and the acoustic proximity of the vowels involved. If the L2 does not influence L3 perception, then subjects from the three L2 groups will have the same results based on L1 transfer only. If L3 vowel perception is influenced by L2 vowels only, all the three groups will exhibit different results, with the exception of the vowels which might be similar in two of the tested L2s. If there is combined cross-linguistic influence, we should see evidence of both L1 and L2 vowels affecting foreign vowel perception.

2. Does selective attention to features (Pajak and Levy 2014) apply to features used in L1 only or also those used in L2 and L3?

Research by Bohn and Best (2012) and Pajak and Levy (2014) suggests that L1 contrastive phonological features affect the way listeners perceive L2 consonants. This study examines whether rounding as a non-contrastive L1 feature or a contrastive L2 feature influences foreign vowel perception. It is hypothesized that learners of L2s with contrastive lip rounding in front vowels should be able to perceive lip rounding in front vowels more consistently than listeners who are not used to distinguishing vowels on the basis of lip rounding.

To guide the answers to these two main questions on the nature of L3/foreign language vowel perception specified above, more detailed phonetic questions related to the study of Dutch vowel perception by Polish learners of English, French and Dutch were also addressed:

3. For Polish learners of L2 English, French or Dutch, do discrimination rates for Dutch vowel contrasts depend on assimilation types, as predicted by PAM?

PAM predictions have previously been attested for non-native, unfamiliar consonants (e.g. Best and Strange 1992, Best, McRoberts, and Sithole 1988) and vowels (Tyler et al. 2014, Faris, Best and Tyler 2016), but the results have been interpreted with reference to the monolingual listeners' L1s. Although differences between monolingual and bilingual listeners certainly influence language acquisition and the effects of L2 speech perception and the initial state in L3 acquisition, it is hypothesized that the cognitive capacities responsible for foreign language speech perception remain the same, so discrimination rates should still depend on assimilation patterns.

It needs to be emphasized that discrimination rates could be higher among L2 listeners than among monolingual naïve listeners because the former are more experienced in discriminating between various non-native contrasts. It is nevertheless hypothesized that these more experienced listeners still rely on assimilation patterns when they discriminate between foreign vowel contrasts.

4. Are there significant differences in discrimination of the eight vowel contrasts, and if so, do relative difficulties differ depending on the L2 or L3? Are discrimination rates proportional to the differences in the acoustic properties of the Dutch vowels in a given contrast?

It was shown that vowel discrimination in the foreign language depends on the L1 vowel inventory (Bohn and Flege 1990 and Best, Bohn, Halle and Faber 2003) and the relationships between the acoustic properties of both L1 and foreign language vowels (Escudero et al. 2012, Escudero and Williams 2012). The four models of third language acquisition assume a vital (potential) role for L2: the Cumulative Enhancement Model by Flynn, Foley, Vinnitskaya (2004), the L2 Status Model by Bardel and Falk (2007, 2012) and Falk and Bardel (2010, 2011), the Typological Primacy Model by Rothman (2011, 2015) and the Linguistic Proximity Model by Westergaard et al. (2017). It is hypothesized, therefore, that non-native discrimination depends on both L1 (here common for all subjects), L2 (English common for all the subjects, but at different proficiency levels) and L3 (French or Dutch only in the case of learners of French and Dutch). This means that if we find differences between learners of English, French and Dutch, these differences will be ascribed to the L3, especially as Antoniuou, Best and Tyler (2013) found no bilingual advantage for unfamiliar contrasts and Antoniuou, Liang and Ettlinger (2015) showed that for difficult contrasts only the subjects with L-specific experience performed better.

In each of the eight tested contrasts /ɪ - ʏ/, /iː - yː/, /ʏ - u/, /ø - u/, /ɛ - ʏ/, /ø - ʏ/, /ɪ - ø/ and /eː - ø/ there is one front or central rounded vowel. It is hypothesized that discrimination rates will be higher for pairs of vowels, in which there is a difference in the height of the vowels, expressed as the difference in F1 measured in Hz in addition to rounding. The hierarchy in the increasing order of differences in F1 appears as follows: /i - y/ (27 Hz) < /ø - ʏ/ (41 Hz) < /ɪ - ʏ/ (59 Hz) < /eː - ø/ (78 Hz) < /ɪ - ø/ (100 Hz) < /ʏ - u/ (145 Hz) < /ø - u/ (186 Hz) < /ɛ - ʏ/ (211 Hz).

5. How are Dutch vowels assimilated to Polish vowel categories?

This is a descriptive question. Theories of second language speech perception speculate that second language sounds are assimilated to the closest native categories. Crucially, the closest native categories need to be established on the basis of perception tests, not phonological proximity. In this study, the tested vowels are actually L3/foreign language vowels (for learners of English and French, and L3 vowels for learners of Dutch), so the configuration of the inventory is more complex. Subjects compared L2/L3/foreign language vowels to L1 vowels, but we assume that also through the "sieve" or filter involving L1, L2 and L3 vowel categories and/or features. Moreover, since Dutch has more than twice as many vowels as Polish does, we can hypothesize that two or more Dutch vowels will be assimilated to one Polish category, but probably with varying goodness ratings.

4.5.2. Method

4.5.2.1. Subjects

The subjects were native Polish advanced/proficient university students of English, French or Dutch, who had completed pronunciation courses in the respective languages, so they were characterized by high metalinguistic competence. None of the subjects had spent more than three months abroad. They learnt their non-native languages in Poland, in a formal school setting. They were taught by Polish teachers in high schools and by both Polish teachers and native speakers at university. There were 26 students of English (hereupon called subjects with L2 English), 12 students of French (called subjects with L3 French) and nine learners of Dutch (called subjects with L3 Dutch).

Participants' profiles are described below on the basis of language history questionnaires, which were filled in by the subjects once they had finished the perception experiment.

The questions referring to the level of competence in a given language asked the subjects to rate it on a six-point scale in accordance with the Common European Framework for Languages (2011). This scale is widely used in the Polish educational system and the subjects seemed confident in estimating their level. A1 is the beginner level, which means that a person can understand and use basic everyday phrases, introduce themselves and interact in a simple way, if the other person speaks slowly and clearly and is ready to help. A2 is the elementary level, at which a learner can understand sentences and frequently used expressions related to immediately relevant matters, such as family, shopping or employment. They can communicate in simple tasks on familiar and routine matters. Learners at the B1 level have intermediate language competence. They can cope with clear standard messages referring to familiar matters frequently encountered in everyday life, including work, school, leisure or travel. They are capable of describing experiences and events, dreams, hopes and ambitions, as well as giving reasons and explanations for their opinions or plans. B2 learners are at the upper intermediate level, which means that they can understand the main points of a complex text on either concrete, abstract or technical topics. They can interact with native speakers fluently and spontaneously. They can express their opinions on a wide range of subjects in a clear and detailed manner, listing advantages and disadvantages of the chosen options. C level denotes proficient users. C1 level learners are advanced learners with effective operational proficiency. They can deal with a wide range of challenging elaborated utterances. They are capable of recognizing implicit meaning. They express ideas fluently and seemingly without effort or being at a loss for words. They are flexible and effective language users in social, academic or professional situations. Their texts, even on complex topics, are clear and well-structured. C2 level users easily understand anything they hear or read. They can coherently summarize information using reconstructed arguments. They express themselves very precisely in complex situations.

Subjects will be called L2 learners of English, and L3 learners of French or Dutch, depending on the language they were majoring in. All subjects in this experiment were familiar with English to some extent, and for most of them English was actually the first non-native language they began learning. Most of the subjects had some familiarity with German, Spanish or French. Nevertheless, their dominant non-native language, in terms of proficiency and frequency of use, was the language they were majoring in at university, i.e. English, French or Dutch.

In terms of the terminology based on the Critical Period Hypothesis (Lenneberg 1967), subjects were L1 Polish listeners, who started having sporadic contact with (mostly non-native) English in early education, had more intensive instruction in English after puberty and started learning another non-native language also after puberty. With reference to the linearity of acquisition, all the subjects in the English and Dutch groups and nine out of 12 subjects in the French group, had contact with English as their first non-native language. According to the terminology used by Hammarberg (2010), Polish was the sole native language for all the subjects, whereas for subjects in the English and Dutch groups and 9 out 12 subjects in the French group, English was the L2. Three subjects in the French group actually had

French as their L2. Finally, all subjects in the English group had rudimentary familiarity with German, Spanish or French. For subjects in the French and Dutch groups, in the course of their studies, their L3 French and Dutch became dominant non-native languages, both in terms of frequency of use and proficiency, although the learners of Dutch and some learners of French still had regular English classes.

Subjects in the English group knew English at the C1 level according to the Common European Framework of Reference for Languages. Since the aim of the test was to examine the influence of an L2 with front rounded vowels, the data from 11 subjects who took part in the experiment, but whose questionnaires revealed knowledge of any language with front rounded vowels above the A2 level (7 subjects knew German at B1 or B2, and 4 subjects knew French at B1 or B2) were discarded. The remaining subjects in the English group also had some basic familiarity with other languages, but as or Gut (2010) claims, the level of L2 must be sufficiently high in order to influence any further languages.

Similarly strict measures were not taken or needed in the case of the French and Dutch groups. Subjects in these two groups knew English above the level of A2. First of all, knowledge of English does not result in familiarity with front rounded vowels, crucial in the present study. Secondly, it would be practically impossible to find Polish students of French or Dutch without any knowledge of English, as it is a compulsory subject in Polish schools. Table 7 summarizes the level of advancement in non-native languages for subjects in the three groups: 26 learners of English, 12 learners of French and nine learners of Dutch. The upper number indicates the number of subjects claiming familiarity with a given language, the information in parentheses indicates the average, self-declared level of proficiency according to the Common European Framework of Reference for Languages, where 1 = A1 or beginner, 2 = A2 or elementary, 3 = B1 or intermediate , 4 = B2 or upper intermediate, 5 = C1 or advanced and 6 = C2 or proficient.

Table 7. The level of advancement in non-native languages

Subjects majoring in:	English	French	Dutch	Spanish	German	Italian	Russian	Norwegian
English (mean age 21.4)	26 (5)	9 (1.67)	–	8 (1.75)	13 (1.38)	1 (3)	3 (2)	1 (1)

French (mean age 22.1)	12 (3.4)	12 (4.9)	–	5 (3.4)	6 (1.33)	2 (2)	–	–
Dutch (mean age 22.8)	9 (4.22)	2 (2)	9 (4.89)	3 (1)	4 (2.5)	–	1 (2)	–

In the light of the above information, the terms *L2 English*, *L3 French* or *L3 Dutch* as used in the present study, should be treated as indicating the non-native language which was dominant for a given subject in terms of proficiency and frequency of use at the time of the study, but no implications as to other non-native languages should be assumed.

4.5.2.2. Stimuli

The recordings took place at the Center for Speech and Language Processing at Adam Mickiewicz University in Poznań, in a studio with an anechoic chamber equipped with digital audio recording gear. A male native speaker of Southern Standard Dutch read carrier sentences with eight Dutch front and central vowels: /ɛ, eː, øː, ɪ, iː, uː, yː, ʏ/ six times. The sentences had the same generic structure as the carrier sentences in Adank et al. (2007). Those sentences are part of the Spoken Dutch Corpus (Corpus Gesproken Nederlands), but as read speech they are not publicly available (one can ask for access to the parts of the corpus with spontaneous speech). The short vowels were embedded in the following carrier sentences: *In sVs en in sVsse zit de V*, /ɪn sVs ən ɪn sVsə zɪt də V/ [In sVs and in sVsse is the V]. The carrier phrases for the long vowels were: *In sVs en in sVze zit de V* /ɪn sVs ən ɪn sVzə zɪt də V/, [In sVs and in sVze is the V]. The CVC contexts were cut out for further processing. Two tokens of each syllable were chosen (minimal differences in length and mean f0). Mean formant values of the vowels used in the experiment are presented in Figure 3.

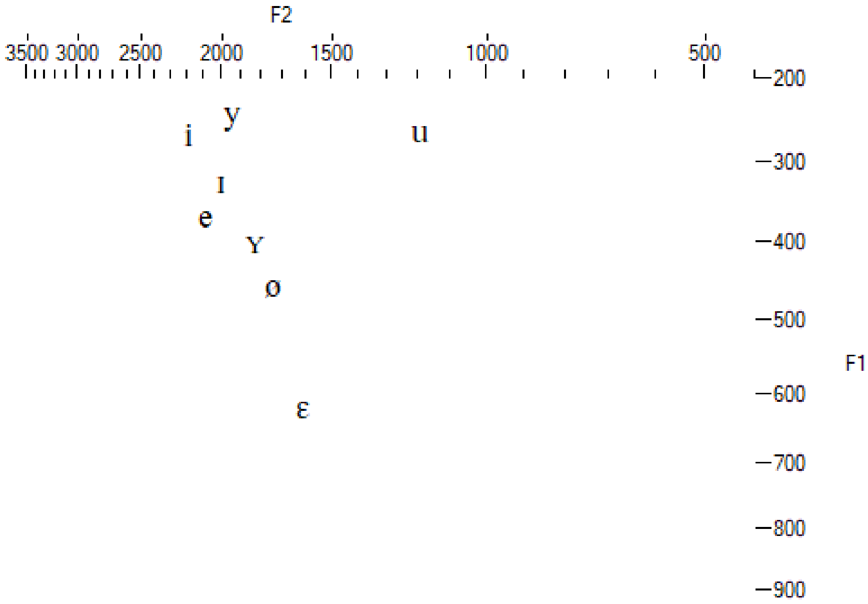

Figure 3. A plot of Dutch vowels used in study two

4.5.2.3. Procedure

The experiment consisted of a discrimination test and an identification test with category goodness ranking of Dutch vowels in terms of Polish vowel labels. First, the subjects performed an AXB discrimination test, in which they listened to 32 AXB triads, in which eight Dutch vowel contrasts were examined: /ɪ - ʏ/, /iː - yː/, /ʏ - uː/, /øː - uː/, /ɛ - ʏ/, /øː - ʏ/, /ɪ - øː/ and /eː - øː/. Each of the eight vowel contrasts was incorporated into a triad, where A was a stimulus from one Dutch vowel category, and B was a stimulus from a contrasting vowel category. Each contrast was presented in each of the four possible triad combinations in an AXB discrimination task: AAB, ABB, BBA and BAA. Each triad combination was repeated three times. The triads with all the tested contrasts were presented randomly. A total of 96 responses were elicited from each subject. In each triad, X represented the same phonological category in Dutch as either A or B, but X was never physically identical to A or B. This step was taken to discourage listeners from focusing on the sheer physical identity of the stimuli. Also, setting a relatively long interstimulus interval (ISI = 1 s, and the intertrial interval = 6 s) was meant to encourage phonological, rather than purely acoustic processing.

In the second part of the experiment, subjects performed an identification task, in which they matched a Dutch auditory stimulus with a Polish vowel label and rated the vowel's goodness in terms of its similarity to the chosen Polish vowel on a Likert scale from 1 (barely matching the Polish vowel) to 5 (well matching). The labels in the task were six Polish orthographic vowel symbols *i, y, e, a, o* and *u* corresponding to the six Polish vowels /i, ɨ, ɛ, a, o, u/. Polish vowel orthography is transparent, so using the orthographic labels was considered to be clear to the subjects. In this task each subject listened to the eight Dutch vowel categories, with two tokens per category, and three repetitions per token, which yields 48 trials. The task was preceded by five warm-up items and afterwards, the stimuli were presented randomly.

4.6. Results

This section presents discrimination results and assimilation with goodness ratings results.

4.6.1. Discrimination results

Discrimination results are the prerequisite to address research questions number three and four. Table 8 shows the mean percentage of correct discrimination, listed in decreasing order, for each Dutch contrast tested in the experiment by the learners from the three groups: L2 English, L3 French and L3 Dutch. Discrimination was excellent for all contrasts (>98%).

Table 8. Mean percentage of correct discrimination of Dutch vowels

Dutch vowel contrast	Language group	% correct responses per each L2 group	% correct responses per contrast
/ɪ - øː/	L2 English	98.72	99.28
	L3 French	100	
	L3 Dutch	100	
/ɛ - ʏ/	L2 English	99.36	99.28
	L3 French	98.61	
	L3 Dutch	100	

/øː - ʏ/	L2 English L3 French L3 Dutch	99.04 99.31 98.96	99.09
/ɪ - ʏ/	L2 English L3 French L3 Dutch	98.40 99.31 100	99.09
/øː - uː/	L2 English L3 French L3 Dutch	98.40 100 98.96	98.91
/ʏ - uː/	L2 English L3 French L3 Dutch	98.56 99.31 98.96	98.82
/eː - øː/	L2 English L3 French L3 Dutch	99.36 97.92 97.92	98.73
/iː - yː/	L2 English L3 French L3 Dutch	98.40 100 97.92	98.73

When individual variation was inspected it turned out that 21 subjects made one mistake, six subjects made two mistakes, two subjects made three mistakes, one subject made 16 mistakes, and the remaining subjects replied correctly to all the 96 triads in the test.

4.6.2. Assimilation and goodness ratings[4]

To address research question number five, Table 9 shows the confusion matrix of assimilation of Dutch vowels / iː, ɪ, eː, ɛ, uː, øː, yː, ʏ/ to Polish vowels /ɪ, ɨ, e, a, o, u/ represented by orthographical letters. The results are listed for each language group separately: Polish learners of L2 English, L3 French and L3 Dutch. The goodness ratings for each Dutch vowel as compared to a Polish vowel are given in parentheses. They are based on a scale that ranges from barely similar (1) to a good match (5).

[4] The preliminary results of assimilation of Dutch vowels to Polish categories presented in this chapter appeared in a much shorter form in conference proceedings in Balas (2017a).

Table 9. Mean per cent assimilation and goodness rating (in parentheses) of the Dutch vowel stimuli to Polish vowel categories

Dutch stimuli	Polish learners of:	Targets: Polish vowel categories					
		i	y (/ɨ/)	e	a	o	u
/iː/	English	99.4 (3.7)	–	–	–	–	0.6 (1)
	French	100 (4.8)	–	–	–	–	–
	Dutch	98.1 (4.4)	–	–	–	–	1.9 (1)
/ɪ/	English	30.4 (3.4)	69 (3.6)	–	–	–	0.6 (1)
	French	5.6 (3.2)	58.3 (3.3)	36.1 (3.8)	–	–	–
	Dutch	50 (3.5)	50 (3.7)	–	–	–	–
/eː/	English	25.8 (3.1)	22.6 (2.7)	48.4 (2.5)	3.2 (1.2)	–	–
	French	1.4 (4)	8.3 (2)	90.3 (3.3)	–	–	–
	Dutch	9.3 (3)	–	90.7 (2.6)	–	–	–
/ɛ/	English	–	–	71.4 (4.3)	28.6 (4)	–	–
	French	–	–	8.3 (4)	91.67 (4.8)	–	–
	Dutch	1.9 (3)	–	88.89 (4.7)	9.26 (4.2)	–	–
/yː/	English	–	45.5 (3.2)	3.6 (3)	–	–	51.95 (2.9)
	French	–	69.4 (3.2)	5.6 (2)	–	8.4 (1.3)	16.7 (3.3)
	Dutch	–	77.8 (3.6)	–	–	–	22.22 (4.5)

/ʏ/	English	2.54 (2)	28.81 (3)	1.69 (2)	–	–	66.95 (3)
	French	7.41 (1.75)	48.2 (3.1)	2.78 (2)	–	0.9 (2)	40.74 (3)
	Dutch	3.7 (2)	51.85 (3)	–	–	–	44.44 (3.1)
/øː/	English	0.6 (2)	18.5 (2.3)	12 (2.1)	–	0.6 (1)	68.2 (2.6)
	French	–	51.4 (2.9)	12.5 (2)	–	15.3 (1.2)	20.8 (2.2)
	Dutch	–	46.3 (2.8)	3.7 (1.5)	–	–	50 (2.3)
/uː/	English	–	–	–	–	–	100 (3.3)
	French	–	0.4 (4)	–	–	–	98.6 (4.2)
	Dutch	–	–	–	–	–	100 (3.9)

The Dutch vowel /iː/ was perceived as Polish /i/ in over 98% of cases by the subjects in the three groups. Goodness ratings suggest that the Polish learners of English and Dutch noticed more discrepancies between the Dutch vowel and the Polish counterpart than the learners of French did.

The Dutch /ɪ/ was perceived as /i/ in 30.4% of cases by the English learners, 50% of cases by the Dutch learners and only 5.6% of cases by the French learners. The Dutch /ɪ/ was assimilated to Polish /ɨ/ in 69% of the cases by the English learners, 50% by the Dutch learners and 58% by the French learners. It is worth noting that 36.1% of the responses to the Dutch /ɪ/ given by the learners of French already pointed to /e/ as a target.

Dutch and French have a high mid vowel /e/ and a low mid vowel /ɛ/. The learners of French categorized the Dutch /eː/ as Polish /e/ (90.3%, goodness rating 3.3) and the Dutch /ɛ/ as Polish /a/ (91.67%, goodness rating 4.8), so this was a Two Category assimilation type. The learners of Dutch categorized both the Dutch /eː/ and /ɛ/ as Polish /e/, but with varied goodness ratings: /e/ with 90.7% and a 2.6 goodness of fit and /ɛ/ with

88.89% and a 4.7 goodness of fit, so this was a Category Goodness assimilation type. Polish and English only have the /e/ vowel in the inventory. For the learners of English, the Dutch /eː/ was an uncategorized clustered vowel, whereas the Dutch /ɛ/ was barely categorized to /e/ with 71.4% and a 4.3 goodness of fit.

The Dutch high front rounded vowel /yː/, which has no counterpart in either Polish or in English, was assimilated by the Polish learners of English to the Polish high back /u/ in 51.95% of the cases, with a 2.9 goodness of fit and to the Polish mid high front centralized /ɨ/ in 45.5% of the cases and a 3.2 goodness of fit. So for the Polish learners of English this vowel was uncategorized at the 70% level, and was barely categorized at the 50% level. If we compare the assimilations to front vowels (49.1) with the assimilations to back vowels (51.95%), we see that the Polish learners of English could not decide whether the Dutch high front rounded vowel better matches a Polish front unrounded counterpart or a back rounded counterpart. The learners of French assimilated /yː/ to the Polish front centralized unrounded /ɨ/ in 69.4% of the cases and if we add the assimilations to the front /e/ (5.6%), three thirds of the assimilations pointed to front vowels. This is similar to the 77.8% assimilations to /ɨ/ by learners of Dutch.

In the discussion section it will be shown how these proportions matter when interpreting the disentangling of rounding from backness by various groups of L2/L3 learners.

The Dutch high mid central vowel /ʏ/ was assimilated by the learners of English to the Polish /u/ in 66.95% of the cases, whereas the assimilations to front vowels amounted to 33.04%. This vowel was an uncategorized clustered vowel for the learners of French, with 58% assimilations to front vowels. The Dutch learners also achieved similar results – 55.55% of assimilations to front vowels, though at the level of 50% this vowel would be categorized to /ɨ/.

The Dutch mid central /øː/ for the learners of English had similar assimilation proportions as /ʏ/ in terms of frontness and backness, but had slightly more assimilations to /e/ than to /ɨ/ and generally lower goodness ratings. The learners of French assimilated /øː/ to front vowels in 63.9% of the cases, whereas the learners of Dutch did so in 50% of the cases. The high back rounded Dutch /uː/ was assimilated to the Polish /u/ by all three groups of learners.

As regards assimilation types of Dutch vowels for Polish learners of the three L2s, the results obtained at the 50% and 70% assimilation thresholds are presented in Table 10.

Table 10. Dutch vowel assimilation types by Polish learners of L2 English, L3 French and L 3 Dutch at 50 and 70% assimilation thresholds

Dutch stimuli	Polish learners of	Assimilation type at 50% assimilation threshold	Assimilation type at 70% assimilation threshold
/iː/	L2 English L3 French L3 Dutch	Categorized Categorized Categorized	Categorized Categorized Categorized
/ɪ/	L2 English L3 French L3 Dutch	Categorized Categorized Clustered	Focalized Clustered Clustered
/eː/	L2 English L3 French L3 Dutch	Focalized Categorized Categorized	Focalized Categorized Categorized
/ɛ/	L2 English L3 French L3 Dutch	Categorized Categorized Categorized	Categorized Categorized Categorized
/yː/	L2 English L3 French L3 Dutch	Categorized Categorized Categorized	Clustered Focalized Focalized
/ʏ/	L2 English L3 French L3 Dutch	Categorized Clustered Categorized	Focalized Clustered Clustered
/øː/	L2 English L3 French L3 Dutch	Categorized Categorized Clustered	Focalized Focalized Clustered
/uː/	L2 English L3 French L3 Dutch	Categorized Categorized Categorized	Categorized Categorized Categorized

4.7. Discussion

This section deals with the discrimination results and assimilation results and moreover discusses the influence of the listeners’ L2 or L3 on Dutch vowel assimilation strategies and summarizes the role of the selective attention to features in L3 and foreign vowel perception.

4.7.1. Discussion of discrimination results

Research question number three asked whether discrimination rates depended on assimilation types, as assumed by PAM. Discrimination rates were ceiling, i.e. above 98%, correct, for all the eight tested vowel contrasts /ɪ - ʏ/, /iː - yː/, /ʏ - uː/, /øː - uː/, /ɛ - ʏ/, /øː - ʏ/, /ɪ - øː/ and /eː - øː/. These results seem to be comparable though higher than the results obtained by Tyler et al. (2014). That study also tested contrasts involving lip rounding: three French contrasts, high-mid front rounded versus low-mid front rounded /dø - dœ/ (96.47% correct discrimination), and high front rounded versus high mid front rounded /sy - sø/ (92.20% correct discrimination), oral versus nasal high-mid back rounded /bo - bõ/ (96.31 % correct discrimination), two Norwegian contrasts, high front unrounded versus outrounded /ki - ky/ (72.76% correct discrimination) and high front unrounded versus in-rounded centralized /ki - kʉ/ (100% correct discrimination), and one Thai contrast, high back unrounded versus high-mid back unrounded /bɯ - bɤ/ (95.51% correct discrimination). Discrimination rates for all the vowel pairs were excellent, > 95%, except for /ki - ky/ for which discrimination was fairly poor (73%). The participants in Tyler et al. (2014) were 13 American English listeners who had not studied French, Norwegian, Thai, or any other language which has the tested vowel contrasts.

In the present study, 21 subjects made one mistake, six subjects made two mistakes, two subjects made 3 mistakes, one subject made 16 mistakes, and the remaining subjects replied correctly to all the 96 triads in the test. When the results were analyzed according to contrast, each contrast had a number of mistakes ranging from 4 to 7 among 552 responses. Since the number of mistakes was marginal and evenly distributed, there was no point in analyzing the mean per cent discrimination scores against individual assimilation types, as was done by Tyler et al. (2014).

The results revealed excellent, ceiling level discrimination rates for all the tested vowel contrasts in all three L2/L3 groups (Table 8). Either the feature [+ rounded] must be easily distinguishable or advanced L2/L3 learners must have acute sensitivity. If one wonders why L2 English learners also find high and mid front and central rounded-unrounded vowel discrimination so easy, we can only ascribe this to experience with the English high central /uː/, which is fronted after a palatal approximant /j/. If this were, however, the case, then American English listeners would have discriminated the Norwegian /i - y/ contrast with more accuracy than 79%

correct (Best, Halle, Bohn and Faber (2003)). Therefore, it seems unlikely that the feature [+rounded] is particularly easy to single out and that one can base discrimination upon it. It is more likely that the advanced L2 learners in this study were really sensitive to phonetic differences.

Technical differences are unlikely to have contributed to ceiling effects, as the same inter-stimulus intervals and inter-trial intervals of 1 s and 6 s respectively were used, as in PAM studies (for example: Best and Strange 1992, Best et al. 1988 and 2001, Tyler et al. 2014). In PAM studies, however, a given contrast is tested within a block, with a randomized order of the four AXB trial types. In the present study, the four AXB trial types were also randomized, but all the contrasts were mixed within one block. This step actually should have presented more of a challenge – subjects did not know which contrast to expect next and could not get used to relying on purely acoustic differences between stimuli in a given contrast so easily. Therefore, this difference between PAM studies and the present study is not likely to have encouraged the ceiling results.

Concluding the answer to research question number three, discrimination rates for all the tested Dutch vowel contrasts in all the three groups were ceiling, so we cannot claim that they depend or they do not depend on assimilation types. It was expected that Cross Boarder contrasts (assimilated as Two Category or Categorized-Uncategorized non-overlapping) would have higher discrimination rates than Category Goodness, and the lowest discrimination rates were expected in the case of Single Category assimilations. The results suggest, however, that subjects did not have difficulty discriminating vowels belonging to any assimilation type. Technical issues are unlikely to have contributed to the excellent discrimination rates. The feature [+ rounded] does not seem to have been a particularly easy feature in other perception studies. The third option ascribes the excellent discrimination rates to the bilingual listeners' sharpened sensitivity to phonetic contrasts in general. This finding would be in line with Antoniou et al. (2015), whose study revealed bilinguals outperforming monolinguals in the case of unmarked features production. In a discrimination task, vowels in a contrast in which A has an unmarked feature combination (as high front unrounded) and B has a marked feature combination (high front rounded) hypothetically should be easier to distinguish than an unknown contrast between vowels which are both marked or both unmarked.

Since discrimination rates for all the contrasts were ceiling, the fourth research question cannot be answered, either.

Generally, more research is needed to shed light on the discrimination of foreign language marked contrasts and unmarked contrasts, discrimination of marked rounded vowels and marked unrounded vowels, as well as comparison of monolingual vs. bilingual foreign language vowel perception and individual differences in such circumstances. Further research should aim at finding explanations for ceiling discrimination rates among bilinguals, despite Categorized – Uncategorized or Uncategorized – Uncategorized overlapping or partially overlapping assimilations.

4.7.2. Discussion of the assimilation results of Dutch vowels to Polish vowel categories by Polish learners of L2 English, L3 French and L3 Dutch.

Research question number five asked about assimilations of Dutch vowels to Polish vowel categories. The Dutch vowel /iː/ is perceived as /i/ by all groups of Polish listeners in over 98% of cases. The goodness ratings are only moderately high, perhaps due to the fact that the Dutch /i/ sound is tenser than the Polish counterpart. There is a discrepancy between how well-matched the subjects perceive the Dutch /i/ to be in terms of the Polish /i/: subjects with L2 English (3.7), L3 Dutch (4.4) and L3 French (4.8). Perhaps the learners of French, not acquainted with the role of vowel length differences, do not pay attention to the longer duration of the vowel in comparison to the Polish standard.

The Dutch /ɪ/ is perceived as /i/ by 30.4% of English learners, 50% of Dutch learners and only 5.6% of French learners, as /ɪ/ by 69% of English learners, 50% of Dutch learners and 58% of French learners. Acoustically, it seems that the Dutch and English /i/ is comparable to the Polish /i/, whereas the Dutch and English /ɪ/ are counterparts of the Polish /ɨ/. Yet both the perception and production of /ɪ/ are notorious problems for Polish learners of English (cf. Bogacka [Balas] 2004, Rojczyk 2010a, Lipińska 2017). The influence of orthography seems to matter here. The letter *i* which in Germanic languages denotes /ɪ/ is read as /i/ in Polish. Most likely, the spelling of pronunciation is transferred onto perception. It is worth noting that in 36.1% of cases, the French learners perceive the Dutch /ɪ/ as Polish /e/. This might stem from the fact that the French /e/ is higher than in other languages, but it also goes in line with the perception of the Dutch /ɛ/ as the Polish /a/, as if front vowels in a chain were perceived as lower.

The goodness ratings of /ɪ/ are around 3.5, which denotes a moderate match of the Dutch /ɪ/ and the chosen Polish categories.

All learners categorized the Dutch /eː/ as Polish /e/, but they assigned relatively low goodness ratings, which is probably due to the duration of this vowel. A considerable number of assimilations of the Dutch /eː/ to Polish vowels /i/ or /ɨ/ were probably caused by the height of the Dutch /eː/. Worthy of note is that the learners of French related the Dutch /eː/ to the Polish /ɨ/, while the learners of Dutch and English, which have tenseness/duration distinctions, related it more eagerly to the Polish counterpart of the Dutch and English high tense vowel, i.e. /i/.

The English and Dutch learners categorized the Dutch /ɛ/ as /e/, while the French learners categorized it as /a/. The phonetic motivation behind it probably is that the French /a/ is front and not completely open and besides there are ongoing changes with respect to the phonemic status of /a/ and /ɑ/ in French. A considerable number of responses by the English learners also pointed to /a/, which might stem from the fact that the three English low vowels /æ, ʌ/ and /ɑː/, which were all perceived as the Polish /a/ (see chapter 3, section 3.9.1. and 3.10.2,), "spread" the boundaries of the category.

The Dutch /uː/ was categorized as the Polish /u/, with mean goodness ratings varying between 3.3 for the English learners and 4.2 for the learners of French. The lowest goodness of fit given by the learners of English might be due to the central quality of the English /uː/ as opposed to the Dutch back /uː/. This result is important as a reference point for discussing the categorization of front rounded vowels.

The Dutch /yː/ is a front rounded vowel, which was perceived as /u/ or /ɨ/ by the learners of English who are not familiar with front rounded vowels and predominantly as /ɨ/ by the learners of Dutch and French who are accustomed to front rounded vowels. Goodness ratings here vary. The learners of English and French assigned lower goodness ratings to /yː/ than to /ɨ/ or /u/. It is worth noting that the learners of Dutch gave higher goodness ratings, which might stem from the fact that the tested language is their L3, and not a foreign language, i.e. they were familiar with the vowels presented.

The results for /ʏ/ were similar to the results for /yː/ -- there was even slightly more balance between /u/ and /ɨ/ categorizations by the learners of Dutch and French, probably because /ʏ/ is more centralized than /y/. For the same reason, the learners of English increasingly chose the /u/ as a target. The learners of English and French similarly evaluated the goodness of fit of the Dutch vowels /yː/ and /ʏ/ to Polish categories, whereas the learners of

Dutch gave higher goodness ratings to /yː/ than to /ʏ/, as if noticing that /yː/ is acoustically much closer to both /i/ and /u/, because they have similar F1 values. This result could be due to the fact that Dutch has more tongue advancement distinctions, so the learners of Dutch might be more accustomed to noticing the features related to the tongue advancement.

When the Polish learners heard the Dutch /øː/, the winning target was /u/ for the English and Dutch speakers, while the French learners preferred /ɨ/. The vowel /ɨ/ was also a common second choice for other learners. Goodness ratings were really low, probably because this Dutch vowel is a front centralized rounded vowel, so it is different from a back rounded /u/ and different from a front unrounded /ɨ/. In acoustic terms, the difference between /y/ and /ʏ/ is larger than the difference between /ʏ/ and /øː/, yet the differences between goodness ratings were larger between the latter two sounds, proving once again that there is no one-to-one relationship between auditory and acoustic properties of sounds. The Dutch and French learners gave higher goodness ratings to /ɨ/ than to /u/ probably because they had been used to hearing non-back vowels being rounded.

On the basis of Table 9, we can see that the feature "rounded" means that a sound containing it was bound to be interpreted by Polish learners as /u/, and it seems that goodness ratings here were proportional to the closeness of the rounded vowel to the Polish /u/. Dutch /uː/ is the closest, followed by /yː/, and followed by /ʏ/ which are much more fronted than /uː/ and then by /øː/ which is as fronted as the /ʏ/, but is also slightly lower than /ʏ/ and much lower than /uː/. In terms of assimilation types, at 70% threshold these were Categorized – Uncategorized, partially overlapping types or, at 50% threshold some of them were Category Goodness types and they would have presented a good testing ground for discrimination performance if only it had not been ceiling.

4.7.3. Discussion of the influence of the listeners' L2 or L3 on their perception of Dutch vowels

Research question one asked whether the listeners' L2 or L3 influences their perception of Dutch/foreign language vowels. In the two experiments, discrimination and categorization with goodness ratings of Dutch vowels was tested. The listeners had a common L1, i.e. Polish, they all knew English to some extent and they had three different dominant non-native languages: L2 English, L3 French and L3 Dutch. The tested language was Dutch.

Discrimination results turned out to be excellent, ceiling rates, in some cases equal to, in other cases higher than the ones reported in the literature for monolingual listeners (cf. Tyler et al. 2014). Therefore we cannot draw conclusions about the language-specific influence of the L2 on L3 vowel discrimination on the basis of the present study. It can only be pointed out that the results might suggest bilingual advantage in foreign vowel discrimination compared to the results obtained by Tyler et al. (2014). Enhanced sensitivity to foreign language vowel contrasts in this study could be ascribed to the cognitive flexibility of advanced L3 learners with high linguistic metacompetence.

The main finding is that identification patterns of Dutch vowels by Polish learners of English, French and Dutch consistently varied across L2 groups. If the subjects all had a common L1, and different dominant L2 and L3s, with different vowel inventories, different sizes and different relationships between the acoustic properties of vowels, we may assume that the dominant non-native languages affected the different results of the Dutch vowel categorization task. The most important finding refers to the perception of the feature [+ rounded]. The Dutch /u/, with its low F2 characteristic for back vowels, was interpreted as /u/ across the board. The Dutch non-back rounded vowels /øː, ʏ/ and /yː/ were identified predominantly as front vowels by learners of languages with front rounded vowels and as back vowels by learners of English who lack experience with the feature [+ rounded] and used with front vowels. With reference to research question number one, these results suggest that experience with L2/L3 categories and features used in new combinations (i.e. front rounded vowels, where the feature [+ rounded] is used in a new context in comparison to the Polish vowel system) exerted an influence on L3/foreign language speech perception.

With regard to models of L3 acquisition, in the case of L3 speech perception of vowels, it seems that we may rule out the hypothesis that it is based on L1 transfer only. If this were the case, then all the categorization results would need to be the same. However, they were not the same. Crucially, we were able to track specific phonetic or phonological explanations for the obtained assimilation patterns (e.g. the shifted perception of front vowels by the learners of French: /ɪ/ was partially perceived as /e/, whereas /ɛ/ was perceived as /a/ or the perception of front and central rounded vowels).

The exclusive influence of L2 on L3 vowel perception would be somewhat difficult to prove in the assimilation task, which specifically asked for a comparison of Dutch vowels with L1 vowels. However, if only L2 categories were considered as targets for L3 vowel assimilation, the learners of French would not have categorized the Dutch /ɪ/ as Polish /ɨ/, since French has neither /ɨ/ nor /ɪ/. The Polish subjects also assigned higher goodness ratings to /ɛ/ than to /e/, showing a preference for the more Polish-like lower /ɛ/, although their L2s have either both vowels (Dutch and French) or only /e/ (English).

Previous phonetic research found L1/L2 to L3 transfer in discrimination, facilitated by experience with the phonemic contrasts present in the tested foreign language and one of the languages known by the subject (Patihis et al. 2015). The present results seem corroborate this claim for categorization. Both the French and Dutch learners, who knew front rounded vowels from their L3s, and the English learners, who had no experience with front rounded vowels, assimilated them to the Polish /ɨ/ and /u/, but the proportions varied depending on L2-specific experience.

The present results also corroborate the conclusion of Antoniou et al. (2015) that in the case of marked features, bilinguals outperform monolinguals in word learning only if they have experience with a given feature. Here, when asked to categorize front rounded vowels, bilinguals managed to disentangle rounding from backness better if they knew a non-native language which included front rounded vowels (i.e. French and Dutch).

On the basis of categorization of Dutch vowels by Polish learners of English, French and Dutch, it seems that the initial stage (Rothman 2011) in L3/foreign language vowel perception is shaped by both the L1 and L2 as well as L3, in agreement with the combined cross-linguistic influence hypothesis (de Angelis 2007). Specifically, L1, L2 and L3 interact with one another both on the phonological level (the number of vowels in the inventory, the contrastive features used in the systems) as well as on the lower phonetic level where acoustic relationships between L1, L2 (cf. Alispahic et al. 2017) and L3 sounds matter. This was evident in different proportions of categorizing a given Dutch vowel to a given Polish target, as well in different goodness ratings dependent on the L2/L3 group. Therefore, the answer to research question number one is that foreign language vowel perception and presumably the initial stage in L3 or further language (L_n) perception exhibit both L1 and language-specific L2 and L3 influences.

4.7.4. Discussion of the role of selective attention to features in L2 and FL

Research question number two asked whether selective attention to features (Pajak and Levy 2014) applies to the features used in the L1 only or perhaps also to the features used in the L2 or L3. The relevant result here is that the Polish learners of French and Dutch, which have front rounded vowels, perceived the Dutch front rounded vowels more often as front than back vowels, in contrast to the Polish learners of English who were more likely to categorize them as a Polish high back /u/. These findings suggest that familiarity with the feature [+ rounded] from the L1 (Polish uses rounding for its back vowels and /w/, but the feature is only secondary, i.e. is not used contrastively/distinctively) does not mean the feature can be easily abstracted and used in a completely different context (front vowels).

There are two possible explanations for the lack of selective attention to the feature [+ rounded] by L1 Polish L2 English listeners. The first explanation would be that the hypothesis about selective attention to features should incorporate markedness. This would mean that a feature known from the L1 is not re-used in perception, if it means opting for an assimilation target with a marked combination of features (front rounded vowels are marked, and are less-frequent in world languages). The other explanation for the lack of, or considerably smaller effects of selective attention to features known from the L1 in a foreign language, would be that the feature in question needs to be contrastive in the L1. This claim would be in line with Bohn and Best's (2012) finding where native German listeners, who do not have /w/ in their L1 inventory, discriminated the /w - j/ contrast better than native English listeners. The authors attributed the enhanced sensitivity to /w - j/ contrast to both the vowel-like properties of word-initial approximants and experience with contrastive lip rounding for front vowels.

The fact that listeners with L3 French or Dutch, where rounding distinguishes front vowels, tended to assimilate front rounded Dutch vowels to Polish front vowels more often than the learners of English supports the second explanation. It seems that for a default monolingual Polish listener the two features [+ rounded] and [+ back] are linked in an implicational hierarchy, in which [+ rounded] implies [+ back]. As an L1 Polish listener you need experience with front unrounded and rounded vowels in your L2, i.e. with an active contrastive feature, to trigger disentangling rounding from backness. Simultaneously, these results mean that selective attention to features can be based on contrastive features from the L2. This gives a positive answer to research question number two.

There remains, however, one more explanation of Dutch vowel perception by Polish learners of various L2/L3s to consider. We cannot rule out the possibility that learners of French and Dutch simply tended to perceive Dutch vowels as front vowels primarily due to familiarity with front rounded vowels. This would mean that the Dutch learners managed to counteract markedness in their L3 and the French learners managed to counteract markedness in a foreign language, just because the tested aspect was similar to what their L3 French experience sensitized them to. According to such a scenario, the above reported results could not be attributed to selective attention to features. The next chapter presents a study designed to resolve the doubt concerning the interpretation of the present results.

Chapter Five

Perception of front rounded and back unrounded vowels

The previous study tested whether there is a language-specific effect of L2 acquisition on foreign vowel perception and whether the acquisition of the L2 with specific vowel features does or does not facilitate the perception of different foreign language vowels with the same feature (cf. Bohn & Best 2012, Escudero & Williams 2012). Specifically, what was tested was how the feature [+ rounded] in various familiar configurations (universal back rounded vowels present in Polish, English, though in a centralized form, French and Dutch) and unfamiliar configurations (front rounded vowels, which are present in Dutch and French, but not in Polish or in English) influences perception of foreign vowels. Dutch front rounded vowels were identified predominantly as front vowels by the learners of French and Dutch and as back vowels by the learners of English. The results suggest that the experience with second language front rounded vowels is enough to trigger disentangling rounding from backness and perceiving the frontness of the vowels, despite their rounding. They also stand in opposition to Gallardo del Puerto (2007), who did not find any bilingual advantage. The previous study, however, did not reveal why the Dutch and French learners succeeded in perceiving the Dutch front rounded vowels as front vowels.

5.1. Hypotheses

To account for the Dutch and French learners' assimilations of front rounded vowels to Polish front vowels we can propose two hypotheses.

- Weak hypothesis: Learners of French and Dutch assimilated front and front centralized rounded vowels to Polish front vowels, because they were simply familiar with these vowels from their L2s:

they managed to perceive the more marked front and front centralized rounded vowels and counteract the universal hierarchy, in which rounding implies backness (see Crothers 1978 and Maddieson 1984). Selective attention to features is limited to contrastive features used in similar contexts in languages known by the listener (cf. the study presented in the previous chapter and Bohn and Best 2012).

- Strong hypothesis: Learners of French and Dutch assimilated front and front rounded vowels to Polish front vowels because there was selective attention to features at play there. Thanks to their experience with front unrounded and rounded vowels in their L3s, learners could learn to actively use the feature [+ rounded]. They should also be able to actively use this feature in a completely new context, not only in other front rounded vowels.

In order to verify the two hypotheses, the present study examines the perception of unfamiliar Turkish vowels: a back unrounded vowel /ɯ/ and a mid front centralized rounded vowel /œ/. In these vowels the feature [+ rounded] is used in a slightly different context than in French or Dutch: /œ/ is slightly lower than /ø/ and /ɯ/ is more marked than front rounded vowels – it is a back vowel devoid of rounding.

A high unrounded Turkish vowel (henceforth HUTV), phonologically speaking, is a back vowel and the evidence for its backness comes from vowel harmony (Kornfilt 1997). Phonetically, its backness was once disputed. Kiliç and Öğüt (2004) explained that, acoustically centralized as HUTV might seem, auditorily it is a back vowel due to the hyperspace effect (Johnson et al. 1993). Also, Ladefoged (1993) observed that removing lip rounding from the back vowel /u/ to produce /ɯ/ raises F2 so that it is seems closer to the center of the vowel formant chart.

The first hypothesis above assumes that the major role in front rounded vowel perception is played by experience with similar front rounded vowels. If this hypothesis is true, the perception of the Turkish vowels, especially of the /ɯ/, will be similar by the Polish learners of English, French and Dutch, because none of the subjects had any experience with high back unrounded vowels. The universal hierarchy of vowel configurations (Crothers 1978) states that the most universal, unmarked system involves front unrounded and back rounded vowels. High front rounded vowels happen to be less frequent and more marked, whereas high back unrounded vowels are the least frequent and the most marked ones. So if sheer experience with a given vowel feature configuration is crucial in foreign vowel perception, then certainly, the French and Dutch learners should differ from the English learners when it comes to front rounded vowel perception, as the latter subjects are not familiar with front rounded vowels. Nevertheless, since a high back unrounded vowel is a more marked feature combination, all three groups of learners are unfamiliar with it, so they should all find such a combination equally challenging in perception. Such a scenario would support a narrow bilingual advantage, along the lines of Patihis et al. (2015). This would mean that the bilingual advantage in selective attention to features is strictly limited to the same or very similar phonemic categories across languages.

If the second, strong hypothesis about the major role of selective attention to features is true, the French and Dutch learners will perceive the Turkish vowels more consistently than the English learners. Due to their experience with front unrounded and rounded vowels in their L2s, the learners of Dutch and French could learn to actively use the feature [+ rounded] to disentangle rounding from backness and faithfully perceive the front rounded vowels. This ability to disentangle rounding from backness will also be used in the case of the Turkish back unrounded vowel and yield more consistent perception by the learners of Dutch and French than by the learners of English. Since bilingualism may increase the ability to focus on relevant stimuli (Bartolotti and Marian 2012), we may also expect that in non-native speech perception those bilinguals whose L2s have front rounded vowels will find it easier to concentrate on or disentangle the phonetic features of rounding and the tongue advancement, because they are already experienced in manipulating them in their L2s, which employ rounding contrastively. Such a scenario would support the idea of a broader bilingual advantage in selective attention to features, not restricted to particularly similar phonemic categories, but based on a broader repertoire of features, and therefore feature combinations at the disposal of L2 learners. The broader bilingual advantage proposed here would not mean that perception becomes more flexible only due to the fact that a person is bilingual, but because more features (from L1 and L2 distinctive feature repertoires) are available for bilinguals to manipulate. This hypothesis offers more possibilities than narrow bilingual advantage based on phonemic categories, but not unlimited advantages due to perceptual flexibility.

5.2. Method[1]

This section presents the third experiment whose aim was to verify the two hypotheses proposed above to account for the results study two.

5.2.1. Stimuli

The recordings took place in a studio with an anechoic chamber equipped with digital audio recording equipment in the Center for Speech and Language Processing at Adam Mickiewicz University in Poznań. The Turkish stimuli were recorded by a male native speaker of Turkish. The speaker read sentences with eight Turkish vowels /i, y, ɯ, u, e, œ, o/ and /a/ in a carrier sentence: *SVs içinde V bulunur* [There is V in sVs]. Only words with /ɯ/ and /œ/ were used for the present experiment. There were also eight Dutch vowels used in the experiment. These were recorded by a native speaker of Northern Standard Dutch. The stimuli were read by the same speaker who had recorded the stimuli in the previous experiment which examined Dutch vowels only, but this time the set of vowels was modified to include: /ɑ, aː, eː, ɛ, ø, u, y/ and /ʏ/. They were recorded in the same carrier phrase and in the same technical conditions as previously. For those categories that were repeated in the current study, two tokens were the same as previously and the third one was a new one (only two tokens of each vowel were used in the previous study, while here in the study three there were three tokens used). All the tokens used in the present study are shown in Figure 4, whereas both their formant values and durations are listed in Table 11.

[1] The preliminary results based on a smaller number of participants and presenting the perception of Turkish vowels only appeared in a much shorter form in Balas (2017b).

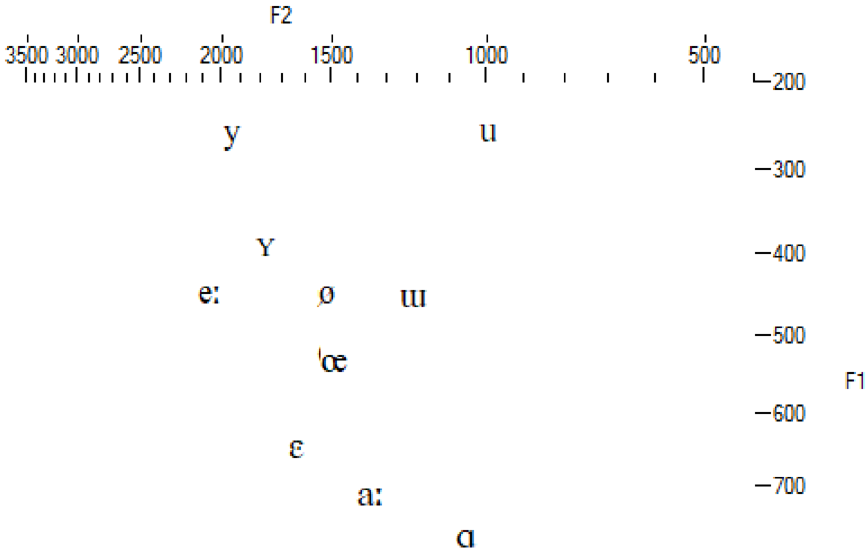

Figure 4. A plot of the Turkish and Dutch vowels used in study three

Table 11. Mean formant values (Hz) and durations (msec.) of the Turkish and Dutch stimuli used in study three

Stimulus	F1	F2	F3	Duration
/ɯ/	438	1260	2934	88
/œ/	492	1556	2616	110
/y:/	240	1941	2289	230
/ʏ/	374	1856	2519	99
/ø:/	422	1575	2333	228
/e:/	427	2124	2797	251
/ɛ/	652	1642	2606	108
/a:/	692	1367	2885	266
/ɑ/	705	1018	2857	123
/u:/	245	1042	2417	204

5.2.2. Subjects

The subjects were native speakers of Polish who major either in English, French or Dutch at Adam Mickiewicz University in Poznań. For all the subjects, English was their second language, i.e. the first non-native language they started learning. As in the case of the experiment on Dutch vowel perception described in the previous chapter, second and third language learners rather than native speakers of English, French or Dutch were chosen, because then subjects from the three groups could categorize foreign language sounds to their L1 categories. In this way a common L1 served as a tertium comparationis for assimilations potentially mediated by the second and third language. There were 27 listeners with L2 English, 22 listeners with L3 French and 25 listeners with L3 Dutch.

Participants' profiles will be described below on the basis of language history questionnaires, which were filled in by the subjects after they finished the perception experiment. The levels of language competence refer to the Common European Framework of Reference for Languages (2011) described in section 4.6.2.1.

The 27 listeners, who were English majors, included 24 females and 3 males. Their mean age was 19.52 (SD = 1,03 range 18 – 23). Their competence in English was advanced, ranging from advanced to proficient (4.76 points on a six-point scale, according to the classification of the Common European Framework of Reference for Languages (2011), based on their declarations and the faculty admission criteria. Their mean age of onset of learning was 6.96 (SD = 2.1). The participants had been learning other languages as their L3 or L3s (i.e. a subsequent foreign language/s after English. Some 20 subjects had been learning German. Their competence was 2.4 on a six-point scale according to the CEFR and the mean age of onset of learning was 9.75. Five participants had been learning French. Their level of advancement equaled 2 on a six-point CEFR scale, and the mean age of onset of learning was 13.83. Five subjects had been learning Spanish. Their level of competence was 2.08, and AOL was 16.67. Three subjects had been learning Russian. Their competence was 2.8 on a six-point scale and the mean age of onset of learning was 12 (SD = 4). Five subjects whose competence in L3 German or French (i.e. the languages with front rounded vowels) exceeded 3, i.e. B1 or intermediate on the CEFR scale, were excluded from the analysis. Only the answers from 27 listeners with minimum competence in L2 English at the level of 4 and the maximum level of competence in L3 French or German not exceeding 3 were analyzed.

The group described here as the learners of French comprised 22 listeners, whose mean age was 21.23. In terms of competence and AOL, their L2 was English. Their competence in English was estimated on average as 4.36

on a six-point scale, with an AOL which equaled 7.82. At the time of the experiment, the subjects had been studying French at university for two years, so in terms of frequency of use, French was their most commonly used non-native language. The subjects declared that their level was 4.03 and the average AOL was 17.36 years. Only in the case of four subjects was the level of competence in French equal to or higher than the level of competence in English. Some 11 subjects knew German at an average level of 2.59 and an AOL of 10.9. In that group one person knew Dutch at the level of C1 and one person knew Swedish at the level of 2.5. Unlike in the case of the learners of English, the responses from subjects who knew German or Dutch at higher levels were also taken into account, as indeed the familiarity with front rounded vowels was expected in this group. Additionally, six subjects knew Spanish at a level of 4.17 and an AOL which was 16.5 and six subjects knew Italian at a level of 2.91 and an AOL which was 18.17.

The group labeled here as the learners of Dutch comprised 25 listeners, whose mean age was 21.58. Like the learners of French, in terms of competence and AOL their L2 was English, which they knew at a level of 4.63 (B2/C1) and with an AOL of 7.46. At the time of the experiment, the subjects had been studying Dutch for two years at university, so in terms of frequency and recency of use, Dutch was their dominant non-native language. They declared their competence in Dutch was at the level of 3.55 and their AOL was 19.54. Only one subject estimated their competence in Dutch as being higher than in English. Additionally, 11 subjects knew German at the A2 level (2.33) and the mean age of onset of learning German was 11.56. Only four subjects declared knowledge of French, which was at the A2 level (2.3), and the mean age of learning was 13.75. Unlike in the case of the learners of English, and similarly to the procedure in the case of the learners of French, high competence in another language with front rounded vowels was not a basis for exclusion from the analysis. Additionally, one subject knew Spanish at the B1 level and two subjects knew Russian at the A2 level.

On the basis of the presented biographical information referring to the language history of the subjects, we can characterize the three groups of subjects.

- In the so called group of learners of English, the subjects were majoring in English and therefore had higher competence in English and were undergoing phonetic training in English. They had limited knowledge of L3s with high rounded vowels, not higher than B1.

- In the so called group of learners of French, the subjects were majoring in French and had undergone phonetic training in French, so the frequency and recency of use point to French as a dominant non-native language, but their overall competence was higher in English, which they had studied for a longer amount of time, beginning in the period before puberty.
- In the so called group of learners of Dutch, the subjects also were majoring in Dutch and had undergone phonetic training in Dutch, so the frequency and recency of use point to Dutch as the dominant non-native language, but their overall competence was higher in English, which they had studied for a longer amount of time, beginning prior to puberty.

5.2.3. Procedure

After the subjects signed informed consent forms, the experiment consisted of an identification test with category goodness ranking of Turkish vowels (and additionally Dutch vowels) in terms of Polish vowels. The subjects performed a keyword identification of the two Turkish vowels /ɯ, œ/ and the Dutch vowels /ɑ, aː, eː, ɛ, ø, u, y/ and /ʏ/ in terms of six Polish vowel categories /i, ɨ, e, a, o, u/ and eight vowel plus glide sequences /ij, ej, aj, uj, ew, aw, iw, uw/. The labels in the experiment were orthographical: *i, y, e, a, o*, *u*, *ej, aj, ij, uj, eł, ał, ił, uł* since Polish vowel orthography is transparent. The tested items were ten vowel categories, three tokens per category, and five repetitions per token, which yielded 150 trials per subject. The stimuli were presented randomly, and were preceded by warm-up items. The subjects clicked on a Polish keyword corresponding to the vowel that they heard in the item and then rated the similarity of the vowel in the auditory stimulus to the vowel in the chosen keyword (1 being barely similar and 7 being a very good fit).

5.3. Results

This section presents assimilation and goodness rating results for the two Turkish vowels and the eight Dutch vowels which were examined in the present study. As there were 10 vowel stimuli in the experiment, 14 responses and goodness ratings, for clarity of presentation, the responses to a given stimulus will be presented in separate tables (only the categorizations above 3% will be shown), whereas the full version of the results is presented in the Appendix. In each table the results for the Polish learners of L2 English, L3 French and L3 Dutch are listed separately. The most frequently chosen identification response per target is boldfaced. Responses which were significantly above the chance level are marked with an *.The goodness ratings were based on a scale that ranged from 1 (barely similar) to 7 (identical).

Perception of the Turkish high back unrounded vowel in Table 12 represents an example of an uncategorized assimilation in the case of both the learners of English and French. The learners of Dutch categorized it as /ɨ/ at the 50% threshold, or, if we assume a 70% threshold, they heard it as an uncategorized clustered vowel, with high and mid front vowels /ɨ/ and /e/ as targets. Listeners from the three groups seemed to categorize this vowel in terms of the Polish /a/ in around 12 to 15% of the cases, although Polish does not share any features with it, apart from the lack of lip rounding. Statistical tests show that assimilation percentages to the Polish /a/ are not significantly higher than the chance level (for English learners t = 1,22 and p-value = 0,12, for Dutch learners t = 0,95, p-value = 0,18, and for French learners t = 1,67 and p-value = 0,055). The learners of Dutch generally seem to interpret it as a mid-high front retracted /ɨ/ or a mid front /e/ – giving a score of 80.51%. The learners of English favored the /ɨ/ vowel, but only with a score of 39.26%. If we add the score of 5.93% for /e/, there seems to be a 45.19% preference for front vowels. The back vowels /u/ and /ɔ/ score 41.23%. In neither case does the value go beyond the 50% categorization threshold. The learners of French perceived the high back unrounded /ɯ/ as a front or back mid vowel, which is in line with the assimilation of other Dutch vowels as lower vowels in comparison to the choices by the learners of English and Dutch. It seems that the learners of English opted for high vowels, either mid-high front centralized /ɨ/ or high back rounded /u/, the learners of French decided to assimilate the Turkish /ɯ/ to mid-high /ɨ/ or mid /ɔ/ vowels, whereas the learners of Dutch assimilated it to the front vowels /ɨ/ and /e/. Concluding, the vowel /ɯ/ was categorized at the 50% threshold to /ɨ/ by the learners of Dutch, otherwise it was an uncategorized clustered assimilation type.

Table 12. Mean per cent assimilation and goodness rating (in parentheses) of the Turkish high back unrounded vowel /ɯ/

Language group	Response keywords: Polish vowel categories				
	y	e	a	o	u
L2 English	**39.26*** (3.57)	5.93 (2,58)	12.59 (2.88)	16.54 (4.06)	24.69* (3.8)
L3 French	29.84* (3.06)	15.56 (2.9)	15.56 (3.29)	**34.60*** (5.45)	3.81 (2.92)
L3 Dutch	**54.36*** (3.01)	26.15* (4.34)	11.54 (3.4)	–	4.87 (4.58)

Table 13 presents the assimilation patterns of another Turkish vowel – a mid front centralized rounded vowel /œ/. Similarly to /ɯ/, this vowel was also uncategorized. The learners of English perceived it as an uncategorized dispersed vowel with /ɨ/, /e/ and /u/ being selected significantly more often than chance. If we concentrate on the tongue advancement, the learners of English perceived the Turkish /œ/ as a front /ɨ/ or /e/ in 55.31% of the cases or as back /ɔ/ or /u/ in 33.33% of the cases. The learners of French perceived /œ/ as mid vowels /e/ or /ɔ/, resulting in an uncategorized clustered assimilation type. The learners of Dutch perceived /œ/ as front centralized /ɨ/ or front /e/, also resulting in an uncategorized clustered assimilation type.

Table 13. Mean per cent assimilation and goodness rating (in parentheses) of the Turkish mid front centralized rounded vowel /œ/

Language group	Response keywords: Polish vowel categories				
	y	e	a	o	u
L2 English	**29.63*** 3.52	25.68* 3.45	7.65 2.97	12.1 3.53	21.23* 3.37

L3 French	16.19 2.82	**43.49*** 3.67	11.11 3.4	24.76* 5.44	3.49 3.18
L3 Dutch	**45.90*** 3.13	43.08* 4.76	3.59 2.71	–	5.13 4.75

When the listeners learning Dutch heard the Dutch /ɑ/ (see Table 14) they chose the Polish /a/ as a target. The Polish learners of English and French, however, split their answers between /a/ and /ɔ/, barely categorizing the vowel to /ɔ/ at the 50% threshold. The Dutch low central /aː/ was assimilated to the Polish /a/, also obtaining rather high goodness ratings (see Table 15). Similarly, the Dutch /ɛ/ was assimilated to Polish /ɛ/ by the learners of English and Dutch (see Table 16). The learners of French, however, also assimilated it to /a/, which is in line with other assimilations in the previous study, where the learners of French perceived Dutch front vowels as lower vowels more than the other groups did. A higher and tenser/longer Dutch /eː/ was assimilated to /ɛ/ and /ɛj/ (see Table 17). For the learners of English it was an uncategorized clustered vowel. For the learners of French it was categorized at the 50% threshold, and uncategorized focalized at the 70% threshold. The learners of Dutch categorized /eː/ to /ɛj/ at the 50% threshold or assimilated it as an uncategorized clustered vowel. Worthy of note is that the learners of French, which does not have length/tenseness distinctions, did not choose /ɛj/ as a target, while it was the main target for the learners of Dutch and the second choice for the learners of English.

Table 14. Mean per cent assimilation and goodness rating (in parentheses) of the Dutch low back centralized /ɑ/

Language group	Response keywords: Polish vowel categories	
	a	o
L2 English	**44.94*** (4.33)	54.07* (6.13)
L3 French	46.98* (5.78)	**52.38*** (5.71)
L3 Dutch	**89*** (5.73)	11.03 (6)

Table 15. Mean per cent categorization and goodness rating (in parentheses) of the Dutch tense front centralized low vowel /aː/

Language group	Response keywords: Polish vowel categories	
	a	o
L2 English	**94.81*** (5.69)	3.46 (4)
L3 French	**98.1*** (5.64)	–
L3 Dutch	**99.45*** (5.1)	–

Table 16. Mean per cent categorization and goodness rating (in parentheses) of the Dutch front mid vowel /ɛ/

Language group	Response keywords: Polish vowel categories	
	e	a
L2 English	**97.78*** (5.41)	–
L3 French	**71.11*** (5.32)	26.03* (4.9)
L3 Dutch	**99.23*** (5.7)	–

Table 17. Mean per cent categorization and goodness rating (in parentheses) of the Dutch tense front high mid vowel /eː/

Language group	Response keywords: Polish vowel categories				
	i	ij (/ij/)	y (/ɨ/)	e	ej (/ej/)
L2 English	6.67 (4.37)	–	10.37 (5.5)	**45.93*** (4.86)	32.10* (4.58)
L3 French	5.4 (2.65)	–	4.44 (6.3)	**68.57*** (4.9)	20.32 (5.53)
L3 Dutch	7.44 (5.9)	3.85 (2.5)	–	21.54* (3)	**66.41*** (4.13)

With reference to Dutch rounded vowels, the high back /uː/ was assimilated to the Polish /u/ by the three groups of learners (see Table 18). The Dutch high front rounded vowel /yː/ was assimilated to /u/ or /u/ plus glide combinations (see Table 19). Worthy of note was that the learners of Dutch assign lower goodness ratings. The Dutch mid high front centralized rounded vowel /ʏ/ was assimilated to the Polish /u/ by the learners of English, but /ɨ/ is also chosen above the chance level (see Table 20). The learners of French assimilated /ʏ/ to /u/ to a lesser degree, chose /ɨ/ more often than learners of English, and assigned lower goodness ratings. The learners of Dutch assigned the lowest goodness ratings to /u/ out of the three groups, and categorized the /ʏ/ to /ɨ/ rather than to /u/. Table 21 shows that the Dutch mid central /øː/ was assimilated as an uncategorized dispersed vowel to /ɨ/, /u/ and /uw/ by the learners of English, but if we add the results for front vowels we reach almost 23% and the results for back vowels add up to 71.12. For the learners of French, /øː/ was an uncategorized clustered vowel. If we sum the responses pointing at front or back vowels, the result is 42.22 for front vowels, and 56.15 for back vowels. The Dutch learners assimilated /øː/ as an uncategorized clustered vowel. In 44.7% of cases /øː/ was assimilated to the front /ɨ/, and in 56.15% of cases to the high back /u/, /uw/ and /uj/.

Table 18. Mean per cent categorization and goodness rating (in parentheses) of the Dutch tense high back rounded vowel /uː/

Language group	Response keywords: Polish vowel categories	
	u	uł (/uw/)
L2 English	**90.37*** (4.89)	8.15 (4.94)
L3 French	**91.43*** (5.2)	7.30 (6.09)
L3 Dutch	**85.38*** (4.4)	12.82 (3.98)

Table 19. Mean per cent categorization and goodness rating (in parentheses) of the Dutch high front rounded vowel /yː/

Language group	Response keywords: Polish vowel categories				
	i	y (/ɨ/)	u	uj (/uj/)	uł (/uw/)
L2 English	–	3.95 (3.63)	**67.9*** (4.09)	12.35 3.9	12.35 4.5
L3 French	–	–	**84.76*** (4.2)	8.57 (4.56)	3.17 (5.4)
L3 Dutch	4.87 (2)	12.05 (3.26)	**59.49*** (2.89)	5.38 (3.14)	17.95* (3.24)

Table 20. Mean per cent categorization and goodness rating (in parentheses) of the Dutch mid high front centralized rounded vowel /ʏ/

Language group	Response keywords: Polish vowel categories		
	y (/ɨ/)	e	u
L2 English	24.69* 4.03	–	**71.85*** 4.56
L3 French	36.83* 3.31	4.13 2.69	**53.65*** 4.27
L3 Dutch	**55.90*** 3.65	–	43.59* 3.8

Table 21. Mean per cent categorization and goodness rating in parentheses of the Dutch tense central high mid rounded vowel /øː/

Language group	Response keywords: Polish vowel categories					
	y (/ɨ/)	e	ej (/ej/)	u	uj (/uj/)	uł (/uw/)
L2 English	15.56* (3.13)	3.21 (2.92)	4.2 (2.59)	**55.56*** (4.5)	–	15.56* (4.49)
L3 French	31.43* (3.53)	10.79 (3.53)	–	**43.49*** (3.85)	1.90 (4.17)	7.30 (4.61)
L3 Dutch	40.77* (3.11)	–	–	**41.79*** (3.19)	1.03 (3.75)	13.33 (4.06)

Generally, the results revealed differences between the three groups of learners. A repeated-measures ANOVA was run on all the Dutch, English and French groups' arcsine-transformed correct identification percentages with Vowel as within-subject factor (10 levels) and L2/L3 as between-subjects factor (3 levels). Since there were no "correct" identification responses, the responses given by the learners of Dutch, as L3 learners of the majority of the tested vowels, were treated as correct targets for the sake of the present analysis. The analysis is supposed to show the differences between assimilation targets for the three groups of learners. There was a main effect of Vowel ($F = 47.688$; $p < 2e\text{-}16$), indicating that some vowels were more difficult than others. There was a main effect of L2/L3 ($F = 17.887$; $p = 2.64e\text{-}08$), which means that there were significant differences between vowel identification between the three L2/L3 groups. The p-value here is really small, but it is slightly larger than for the stimulus, which indicates that the stimulus influenced the results more than the L2/L3. There was a significant Vowel x L2/L3 interaction ($F = 4.773$; $p = 3.81e\text{-}10$) (more than for the Vowel, and less than for the L2/L3), which suggests that the three groups differed in their identification responses for some Dutch vowels.

5.4. Discussion of assimilation results and the comparison with study two

The major issues to be solved now within PAM are the following: to evaluate discrimination predictions for uncategorized phones and the influence of perceptual overlap on discrimination rates; *new* category formation and replacement of the use of an arbitrary categorization threshold (Faris et al. 2016). The present study may not contribute to these aims, but it uses the three uncategorized assimilation types distinguished by Faris et al. (2016) to trace the perception of Dutch and Turkish vowels by the three groups of Polish learners of Dutch, French and English.

This section discusses the assimilation results of the present study on Turkish and Dutch vowel perception by Polish learners of English, French and Dutch, and also compares them, where relevant, to assimilation responses in the previous study reported in sections 4.6.2. and 4.7.2. In the previous study, perception of the following eight Dutch vowels was tested: /iː, ɪ, eː, ɛ, yː, ʏ, øː/ and /uː/. In the present study, the main focus is on

Turkish vowels /ɯ/ and /œ/, but some Dutch vowels were retested or introduced to the study as well. Although participants in both studies shared the main characteristics (L1, L2, L3 and age group), some of the vowels were retested, because the subjects who took part in this study were not the same individuals who participated in the Dutch vowel perception study. For the sake of comparison between perception of the unmarked high back rounded vowels and marked high front and central rounded vowels, the assimilation of /uː, yː, ʏ/ and /øː/ was retested. In order to confirm the /ɛ/ identification by the French learners as Polish /a/, the Dutch /eː/ and /ɛ/ vowels were retested and the Dutch /ɑ/ and /aː/ vowels were also included.

With regard to the assimilation patterns of /eː, ɛ, aː/ and /ɑ/, the results showed language-specific differences. High-mid front /eː/ was categorized as Polish /e/ or /ej/, and the latter vowel plus glide combination was chosen above the chance level only by the learners of English and Dutch, who have experience with long/tense vowels. The lower and shorter/laxer /ɛ/ was categorized as Polish /e/ by all three groups of subjects, but the learners of French also gave a significantly higher than chance number of /a/ responses – 26.03% (t = 2,54 and p-value = 0,0097). This is in line with, but not identical to, the categorizations in the previous study, where even more preference for /a/ was shown: the learners of English chose it in 28.6% of cases and the learners of French in 91.67% of cases. The difference in categorization ratios in the present and previous study can be interpreted as stemming from the lower language advancement and experience of the participants in the present study.

The Dutch mid low central /aː/ was categorized as the Polish /a/. The Dutch low back /ɑ/ was categorized as the Polish /a/ by the learners of Dutch (89%), whereas the learners of English and French chose /ɔ/ (54.07% by the learners of English and 52.38% by the learners of French) and /a/ (44.94% by the learners of English and 46.98% by the learners of French) as targets, probably because both English and French have mid low back rounded vowels in their repertoires.

As for the rounded vowels, the Dutch /uː/ was perceived as the Polish /u/. It is worth noting that the Polish learners of French assigned higher goodness ratings to the tense/long vowel, but it seems again, like in the case of /eː/, that they may have not noticed the additional tenseness or length of the Dutch vowel in comparison to the Polish counterpart.

The Dutch /yː/ was generally perceived as /u/, or /u/ plus glide. This stands in opposition to the previous study, in which the majority of responses given by the Dutch and French learners and a significant number of responses by English learners pointed to the front vowel /ɨ/. There were three differences between the previous and the present study: the language proficiency of the subjects, more response options in the present study, including the vowel plus glide sequences, and more tokens employed and reused in the present study. One or more of these factors must be responsible for the difference in the results. Since the tokens were almost identical, it seems that it was either the level of L2 or response options that influenced the results.

In the case of a front centralized high mid rounded /ʏ/, an identification pattern can be observed: the learners of English categorized the vowel as a Polish front centralized unrounded /ɨ/ only in 24.69% of cases, the French learners in 36.83% of cases, and the learners of Dutch in the majority of cases (55.90%). The more experience the subjects had with front rounded vowels, the more likely they were to choose a front vowel as an identification target for the rounded mid-high front centralized /ʏ/. These results do not in fact differ from those obtained in the previous study (see section 4.6.2.).

For the Dutch mid central rounded /øː/ responses were split between front and back vowels. The learners of English were more likely to choose back vowels as targets, and they differed in the percentage of /u/ responses from both the French (z = 3,212 and p-value = 0,00132) and Dutch learners (z = -3,88 and p-value = 0,00010).

In the light of the two hypotheses on selective attention to features proposed above, it is difficult to interpret the data. The weak hypothesis assumed that L2/L3 learners can only find it easy to perceive a given foreign language sound if it is very similar to one of the sounds in their L1, L2 or L3 repertoire. If the weak hypothesis was true, all groups should find categorizing the /ɯ/ challenging. The strong hypothesis claimed that selective attention to features should enable those who are familiar with high front rounded vowels (i.e. who are used to employing the feature [+ rounded] contrastively) to consistently perceive the Turkish high back unrounded vowel. In the case of the initial results of this study, for a smaller number of participants (17 learners of English, 10 learners of French and 25 learners of Dutch) reported by Balas (2017b), it seemed that the learners of French and Dutch were able to disentangle rounding from backness and categorize /ɯ/ at the 50% threshold (the learners of French to /ɔ/, and the learners of Dutch to /ɨ/), whereas the learners of English chose four Polish vowel categories at rates between 20 and 27%. It seemed that the learners of English did not know how to react to such a combinations of features: [+ high], [+ back], [- rounded]. At the same time, the learners of Dutch and French, who were accustomed to manipulating lip rounding in the case of their L3 front vowels, found strategies to consistently perceive the Turkish /ɯ/ as /ɔ/ (the learners of French) and as /ɨ/ (the learners of Dutch).Those initial results seemed to support the strong hypothesis. After adding data from more participants, the results became less clear. The learners of English chose two vowels /ɨ/ and /u/ above the chance level, whereas the other two targets were chosen not significantly more often than the chance level, one vowel below the chance level. The learners of English decided to choose high and mid-high vowels as targets. The learners of French, who initially seemed to have categorized /ɯ/ as /ɔ/, now assimilated it below the categorization threshold to /ɔ/ – 34.60% and to /ɨ/ – 29.84%, as if they had opted to categorize /ɯ/ to mid-high and mid vowels. The learners of Dutch assimilated /ɯ/ to the Polish /ɨ/ above the 50% threshold, and also pointed to /e/ above the chance level, as if they had decided to choose front mid-high and mid vowels. Perceiving a given vowel as uncategorized is nothing wrong – after all it should be easier to create a new category for a non-native sound, if that sound is not assimilated to one of the L1 sounds. When listeners do not categorize a non-native sound, they signal that they notice differences between a given non-native sound and the L1 categories. The point is that in the preliminary report by Balas (2017) it seemed that the learners of English did not know how to react to /ɯ/, whereas the learners of French and Dutch had their strategies to disentangle the lack of rounding from frontness.

The results of the previous study from chapter four suggested that learners of French and Dutch who had experience with high and mid front vowels, were more likely to interpret them as front as opposed to back vowels in comparison with learners of English who did not have experience with contrasts based on rounding. This was interpreted as the in/ability on the part of English learners and French or Dutch learners respectively to disentangle rounding from backness. Once a non-native language has a distinction based on rounding in the front region, learners realize that a rounded vowel does not necessarily signal /u/. They could ignore rounding and concentrate on perceiving the height and advancement of a vowel. It was assumed that in the same manner the learners of French and Dutch would be able to ignore the lack of rounding and decode /ɯ/ as a high back vowel. The results have, however, not supported the strong hypothesis on selective attention to the feature [+ rounded]. The only learners who actually, to a significant degree, chose the Polish /u/ as a target for /ɯ/ were the learners of English, who do not have any experience with contrastive rounding. Their L2 English only has a high front spread /iː/ and high central slightly rounded /uː/, which is a back vowel only before a velarized /l/ or becomes front after /j/. This may mean that their category boundary for /u/ is generally spread beyond the traditional Polish high back region and covers the high central and fronted areas. If we also take into account the results of the study in chapter three, we should assume that the English mid-high central slightly rounded /ʊ/ is also subsumed under the Polish /u/ category, probably further stretching its boundaries to encompass the mid-high central region. The present results of /ɯ/ categorizations do not seem to support the strong hypothesis on selective attention to features. They rather seem to suggest that the subjects assimilated the tested vowels to any category that under the influence of L1, L2 and the L3s known by the subjects seemed in some respect similar to the tested item.

Concluding, it seems that in non-native vowel perception the known categories from L1, L2 or L3 prevail. There is only some evidence for the influence of selective attention to features in the case of more marked feature combinations, unknown to the listeners from any language they are familiar with. On the basis of the two studies it seems that consistent perception of marked feature combinations is more easily attested in the case of familiar feature combinations – as was the case with high front rounded and mid front rounded vowels, with which the learners of Dutch and French were familiar. More marked feature combinations than the ones found in L1, L2 or L3 do not easily subscribe to selective attention to features. On the other hand, it cannot be claimed that the weak hypothesis is confirmed, as it assumed that /ɯ/ should be equally challenging for all groups of learners. It was not. The Dutch learners even managed to categorize it. These results seem to suggest that fine-tuned phonetic categories shape perception. All the vowels in the languages one knows, shape, spread and modify the vowel phonetic categories. These unique category configurations then influence the way non-native or foreign vowels are perceived. Perhaps the present results offer some support for exemplar theories.

The results support the narrow-based L1/L2/L3 to L3/FL transfer: if the subjects knew an L3 with a similar sound to the FL sound tested, they were able to decompose its features more efficiently and assimilate it more consistently to L1 categories, even in the case of marked feature combinations. In the case of a highly marked unfamiliar feature combination, as with the Turkish /ɯ/, the advantage seems to be less clear, not pointing directly at a fully-fledged selective attention to features. Those features which are more marked than the ones familiar from the L1/L2/L3.

Chapter Six

General discussion

> "But for our science, whether we call it phonology or phonetics,
> it is what is meant that is the given,
> and how it is expressed that is the puzzle
> – why we speak with such unconscious and inadvertent noise,
> and why that noise, however massively variable,
> gradually reveals such lawlike regularities."
>
> Stampe and Donegan (2009: 26)

In the light of previous abundant evidence for the crucial role of categories which act as magnets, the role of individual features and selective attention to features could only be supplementary. We have, nevertheless, found numerous examples of the isolation of features and use of the known ones in a facilitative way in non-native speech perception.

6.1. Proofs of isolation of features from L2 categories

Any language has a basic set of vowels and uses basic features; it is not possible to test a given feature in complete isolation from other features. Any discussion of selective attention to features therefore needs to be based on the features abstracted or isolated by listeners from category specifications.

6.1.1. Duration

Scenarios predicting the role of duration in L2, L3 and foreign language perception by Polish listeners were, on the one hand, based on the desensitization hypothesis (Bohn 1995) as well as studies showing the reliance of Polish learners of English on duration cues: Bogacka (2004) for the /iː - ɪ/ contrast, Rojczyk (2010b and 2011) for the /æ - ʌ/ contrast, Porzuczek (2010) for the improvement in the weak form *to* as a function of experience, Rojczyk (2013) for the distinction between the verb and noun *record*. On the other hand there were studies pointing to challenges posed by duration to Polish learners of English in the case of differentiating between long and short vowels (Waniek-Klimczak 2005) and pre-fortis clipping (Waniek-Klimczak 2005 and Szpyra-Kozłowska et al. 2014), which would support the feature hypothesis (McAllister et al. 2002).

The results of study one did not reveal particularly strong attention to duration. Discrimination of English vowel contrasts did not rely on duration at all. Out of six excellently distinguished contrasts, only two (or three if we assume longer duration of /æ/) had duration differences: /iː - ɪ/, /e - æ/, /e - ʌ/, /e - ɪ/, /ʌ - ɒ/, /ɔː - ɒ/. All these contrasts were primarily distinguished by tongue height. The contrasts with lower discrimination rates such as /e - ɜː/, /æ - ʌ/, /ʌ - ɑː/ and /uː - ʊ/ all had duration differences. As for (dis-)similarity ratings for English vowel pairs, the least similar vowel pairs, that is /e - æ/, /e - ɪ/ and /e - ʌ/, do not have duration/tenseness distinctions. The results suggest that listeners did not pay attention to duration, but only to assimilations to categories and height differences. The results of goodness ratings of English vowels as assimilated to Polish categories suggest that duration differences together with spectral differences were responsible for lower goodness ratings for free as opposed to checked vowels. Identification results reveal that free vowels /iː/, /ɔː/ and /uː/ were well identified. Nevertheless, /ɜː/, whose mid-central quality is new for L1 Polish listeners and /ɑː/ which is assimilated to the Polish /a/ together with two other English vowels, were not so easily identifiable. Worthy of note here is the fact that among the poorly identified low vowels which were assimilated to the Polish /a/, the long/free /ɑː/ was the easiest one to identify. Duration seemed to mediate misidentifications in that free vowels were identified as other free vowels, whereas checked vowels were identified as checked vowels: /ɜː/ was interpreted as /ɑː/, /ʌ/ was interpreted as /æ/ or /ɒ/ and rarely as /ɑː/, /ɑː/ was interpreted as the longest checked vowel /æ/ or as /ɔː/, but never as /ɒ/ or rarely as /ʌ/, /ɔː/ was interpreted as /uː/ much more often than as /ʊ/, which is closer to /ɔː/ in terms of F1.

Concluding, in study one the role of duration was of secondary importance: (1) duration coupled with spectral differences caused lower goodness ratings of the long/free vowels, (2) free/long vowels were generally better identified than short/checked vowels unless special spectral characteristics or Single Category assimilation were involved, (3) free/long vowels were substituted for other free/long vowels, whereas short vowels were substituted for short vowels. Duration differences did not, however, boost discrimination rates or (dis-)similarity ratings.

6.1.2. Vowel height

Universally, vowel height distinctions are equal to or more numerous than backness distinctions (Crothers 1978). The results of the experiments presented here also showed that vowel height distinctions are relatively easily perceivable. In the task where the subjects assimilated English vowels to Polish categories, differences between the height of an English and Polish vowel were easy to notice and such assimilations received lower goodness ratings than in the case of backness distinctions. We observed relatively high goodness ratings, i.e. above 4.6 for the English /iː/, in phonotactically permissible environments for /ɪ/, for /e, æ/ and /ɒ/, whose height characteristics are similar to those of Polish vowels. In the case of English vowels which differ more in height from Polish vowels, goodness ratings were below 4.5: /ʌ, ɔː/ and /ʊ/ (admittedly, /ɔː/ differs from the Polish counterpart both in terms of vowel height and duration) . Other vowels which received lower goodness ratings were /ɜː, ɑː/ and /uː/. They may not be considerably different from Polish vowels in terms of vowel height, but they differ from Polish vowels in terms of two other features, namely backness and duration, so there are other reasons for lower goodness ratings here.

In fact, it was not combinations of two known features which eased perception, as the hypothesis proposed. Two remarkably different features yielded lower goodness ratings. Lower goodness ratings mean that equivalence classification did not hinder isolation of the different features of a non-native sound in comparison to the native one. They suggest that a learner is on their way to forming a new category for a given sound, which is the most welcome scenario.

In the English vowel identification task, the subjects confused vowels in the horizontal plane, but a very marginal percentage of incorrect categorizations had targeted vowels of a considerably different height than the stimulus. There were 11 English categories to choose from, so the chance level was 9%. The /iː/ vowel was not mistaken for /ɪ/, the /ɪ/ was in 12.1% of the cases misinterpreted as /iː/, but this happened in velar contexts, where

Polish has a phonotactic restriction on the occurrence of /ɪ/. The English /e/ was not mistaken for /æ/, and vice versa (the stimulus for /æ/ was a modern British English mid-low front centralized vowel). Low vowels were mainly confused with one another, vowel length mediating confusions with vowels of different height: /ʌ/ was mistaken for /ɒ/, whereas /ɑː/ was mistaken for /ɔː/. The high-mid central /ʊ/ was sometimes confused with the mid-high front centralized /ɪ/, but not really with higher /uː/. The results suggest that a confusable feature is backness, but not height.

Vowel height was also a crucial feature in the English vowel discrimination task: the excellently discriminated contrasts, such as /iː - ɪ/, /e - æ/, /e - ʌ/, /e - ɪ/ and /ʌ - ɒ/ (admittedly, there is a relatively small height difference here, perhaps the very good discrimination performance should be subscribed to the difference in lip rounding), and /ɔː - ɒ/, involved a difference in the height of vowels. The poorly discriminated contrasts primarily involved backness distinctions such as, for example, /e - ɜː/ and they were Single Category assimilation types: /æ - ʌ/ (there is a height difference between these two vowels, but probably assimilation of both to the Polish /a/ with similar goodness ratings suggests that equivalence classification is at play here), /ɑː - ʌ/ and /uː - ʊ/. Vowel height differences seemed easy to spot in the English vowel discrimination task, as opposed to vowel backness. The two exceptions were explained by resorting to another differentiating feature or equivalence classification.

As the results of the discrimination task generally matched the results of the (dis-)similarity rating task, there is also a link between the height differences in the examined vowel pairs and the (dis-)similarity ratings. The three Single Category contrasts (/æ - ʌ/, /ʌ - ɑː/ and /uː - ʊ/), the latter two without a considerable height difference, were the ones judged to be most similar. The contrast /ɒ - ɔː/, rated as somewhat similar, actually involves a height difference, but the vowels also have two features in common: backness and rounding. The /ʌ - ɒ/ contrast was also rated as somewhat similar, yet it does not involve much of a height difference, just a relatively small difference in backness and rounding. Vowel contrasts considered to be dissimilar included the following height differences: /iː - ɪ/, /ɒ-ɑː/, /e - ʌ/, /ʊ - ɜː/, /e - ɪ/ and /e - æ/.

The discrimination of Dutch vowels in study two revealed ceiling results, but actually all the stimuli differed in vowel height, and also either in lip rounding or backness. Perhaps this is the reason for ceiling results – two differentiating features, in addition to the bilingual factor which might have also enhanced discrimination. In the Dutch vowel assimilation task, the more distant a given Dutch front rounded vowel was from Polish /ɨ/, the lower the goodness ratings it received. The remaining experiments in study two and study three were designed primarily to test the effects of experience with lip rounding in various contexts. Interpreting their results is undoable without a three-way comparison of formant values for vowels in all the languages discussed.

Concluding, it seems that vowel height is a very salient feature. The differences in height between English and Polish vowels caused lower goodness ratings. In the English vowel identification task, vowels of various heights were not confused with one another. Vowel height differences were linked with higher discrimination rates and dissimilarity ratings.

6.1.3. Tongue advancement

Since backness distinctions are never more numerous than height distinctions (Crothers 1978), we might expect that the backness feature is less robust than height. In the task where the subjects assimilated English vowels to Polish categories, differences between the backness of English and Polish vowels were not too frequent, and they were also accompanied by differences in other features. For example, the English /ɜː/ is a mid-central vowel, whereas the Polish /e/ is front central, the English /ɑː/ is much more retracted than the Polish /a/, and the English /uː/ is a high central vowel with diphthongization and less lip rounding than the Polish /u/, but there is also a difference in duration – probably the differences in both cues, backness and length, contribute to the rather low goodness ratings. At this point, there is no explanation as to why the English /ɜː/ scores so much lower than /ɑː/ and /uː/.

In the English vowel identification task, mainly assimilation types and duration mediated the answers, but within Single Category assimilations we can observe cases of misidentifications in which backness differences were overridden, which did not happen in the case of duration or height.

Within the English vowel discrimination results, we can clearly see that contrasts with height differences were the best discriminated, contrasts with backness distinctions were the most poorly discriminated, and differences in duration did not matter. Also, in the (dis-)similarity rating task, contrasts with backness distinctions were deemed more similar than contrasts with height distinctions. The universal tendency of vowel systems to have more tongue height than backness distinctions is confirmed here: the tongue height differences were more discernible than tongue advancement differences, and vowels involving tongue advancement differences were perceived as more similar than those which differ in tongue height.

6.1.4. Lip rounding

Rounding is a feature which occurs in Polish in an unmarked combination only – in back rounded vowels /u/ and /ɔ/. The English /uː/ is generally a high central vowel with fronted allophones after /j/ and retracted allophones before the velarized [ɫ]. Neither in Polish nor in English is rounding a contrastive feature. French has three front rounded vowels /y, ø/ and /œ/, whereas Dutch has four: /y, ʏ, ø/ and /œ/. In French and Dutch, front unrounded vowels contrast with front rounded vowels. The front rounded vowels are typologically rare and they are considered to be marked. The Turkish high back unrounded /ɯ/ is even more marked. Testing the perception of the above-mentioned English vowels by learners of English, and Dutch and Turkish vowels by Polish learners of English, Dutch and French allowed for observations related to rounding and markedness.

In study two, in which assimilation of Dutch vowels to Polish categories was tested, the results showed that Polish learners of French and Dutch were more likely to assimilate Dutch front rounded vowels to the Polish front /ɨ/ rather than to the back /u/ (the lower and the more central a given Dutch vowel was, the more likely it was to be interpreted as /u/). Such results were taken to suggest that either experience with front rounded vowels in their L3 allowed them to notice that these vowels were front rather than back, or they simply learnt to isolate the feature [+ rounded] from backness and they were able to manipulate it. In study three, where the subjects were asked to assimilate the Turkish high back unrounded /ɯ/ to Polish categories, it seemed that neither group was able to disentangle the lack of rounding from front vowels, which refuted the strong hypothesis regarding selective attention to features in more marked, unknown combinations. One of the more interesting findings here was that actually the learners of English were the only ones to perceive the Turkish /ɯ/ as a high back vowel, which was ascribed to the broad /u/ category which they might have formed to cover the Polish high back rounded /u/ and the English high central slightly rounded /uː/, together with its fronted and retracted allophones, and the English mid-high central slightly rounded /ʊ/.

In the English vowel assimilation and identification tasks in study two, rounded vowels were not assimilated to or categorized as unrounded vowels and vice versa. Unfortunately, it is not possible to say how much slight lip rounding in the English /uː/ and /ʊ/ contributed to lower goodness ratings in the assimilation task, because in both cases there are co-existing features which were different: centralization of both vowels, duration of the /uː/ and lowering of the /ʊ/. Certainly, discriminating between the two central rounded vowels was difficult for the subjects, which is in line with assimilation of these phones as a Single Category. The subjects noticed dissimilarity between /ʊ - ɜ/ and /ɒ - ɑː/, but they considered /ʌ - ɒ/ to be rather similar, so maybe it is the difference in duration and rounding that results in lower goodness ratings (duration differences on their own did not promote dissimilarity).

Polish listeners, who are not familiar with contrastive lip rounding from their L1, showed that they were able to perceive rounding in front rounded vowels if their L3 has front rounded vowels, and that they were able to adjust the Polish /u/ category to encompass the English high central and slightly rounded /uː/ as well as the mid-high central slightly rounded /ʊ/. Assimilating the marked Turkish high back unrounded vowel was more complex than initially assumed.

6.1.5. Hierarchy of features

The proposed hierarchy of readily perceivable features by Polish listeners, learners of L2 English, L3 French or L3 Dutch, places vowel height at the top as the most robust cue, as vowel height is not prone to being misperceived and height differences are easily noticeable. Lip rounding comes second as it determines whether vowels are perceived as front or back, disentangling the two is difficult, and rounded vowels are not mistaken for the unrounded ones when the unmarked combinations are tested (as in study one). Duration is not as reliable a cue as rounding: it does not influence discrimination or (dis-)similarity ratings. Prolonged duration, nevertheless, together with spectral differences, yielded lower goodness ratings. Free/long vowels were generally better identified than short/checked vowels, except for the special case of /ɜː/ and Single Category assimilations of low vowels. In the identification task, free/long vowels were misidentified as other long vowels, whereas checked/short vowels were misidentified as other short ones. Vowel backness is the least robust of the cues analyzed here: it was often misinterpreted in the Dutch and Turkish vowel perception tasks; in the English vowel discrimination task, contrasts differing in backness were the worst discriminated ones; and in the (dis-)similarity rating task, contrasts considered to be the most similar were the ones differing in backness. In fact, the three cases of Single Category assimilation types in the English vowel assimilation task involved vowels primarily differing in backness. Hence, the proposed hierarchy of perceivable features is as follows: vowel backness < duration < lip rounding < vowel height.

6.2. Conclusions regarding the theories of speech perception

The phonological models of second language speech perception, the Speech Learning Model (Flege 1995), the Perceptual Assimilation Model (Best 1995, Best and Tyler 2007), the Native Language Magnet Model (Kuhl 1994, 2000a; Kuhl et al. 2008) and the Second Language Linguistic Perception Model (Escudero and Boersma 2004, Escudero 2005 and 2009) have been backed by numerous studies supporting the primary role of categories in non-native speech perception. The studies presented here were not meant to refute the role of categories, but to provide evidence for the role of selective attention to features. This has been achieved by showing that in second, third and foreign language perception, distinctions based on features or feature parameters known from the L1 or previously acquired languages are easier to perceive (height or lip rounding), while those which are not used contrastively are more challenging to perceive and their effect surfaces in a restricted number of contexts (duration). When making predictions for non-native perception, both universally salient features, such as height, as well as the features known/unknown to subjects with a specific language combination could be taken into account to make the predictions more precise.

Language specific L1, L2 and L3 influences could be observed in the results of studies two and three. The results also indirectly suggested bilingual advantage in discrimination performance. The models of third language speech should take into account that experience with the L2 and L3, in addition to the L1, influences the perception of L3 or foreign sounds.

6.3. Directions for further research

The studies in this book concentrated on group results. Nevertheless, monitoring individual variation in future studies on selective attention to features would be worthwhile in the light of recent interest in how individuals differ in their perception of speech. The methods for tackling individual variation in non-native speech perception have been shown, for example, in Tyler et al. (2014) and Mayr and Escudero (2010). Tyler et al. (2014) split subjects' results by assimilation type, rather than by contrast. Mayr and Escudero (2010) showed that English learners of German follow different paths in the developmental perception of German vowels. Their assimilations were highly diverse, yet very systematic. Determining whether the same kinds of regularities apply in L3 and foreign language perception would allow us to account for at least a portion of the variability in post-L2 perception. In the case of the present study one, analyzing individual assimilations would allow for determination of individual developmental paths for advanced learners in a formal instruction setting. Since the present results in stage one and two did not differ in many respects, spreading the time window would be a reasonable idea, too. In studies two and three, the central question would be how much individual variability there is when sounds are assimilated as uncategorized. This would allow for the extension of Faris et al.'s (2016) work on uncategorized phones and the hierarchy of discrimination predictions.

The ceiling results in discrimination which were obtained in study two also need further explanation. Was the L3 and foreign vowel perception enhanced by the presence of two features differentiating each contrast (i.e. lip rounding and height or lip rounding and duration)? Or is the ease of perception attributable to the linguistic experience of highly proficient L2 and L3 users? If so, how much experience is needed to reach the level of bilingual advantage for L3 and foreign language vowel discrimination? As vowels are generally less categorically perceived and their discrimination is easier, how much experience would then be needed to reach bilingual advantage in consonant perceptions?

Since, in fact, only a little progress in perception has been captured in study one, it would be worth checking whether subjects can develop their perception more as a result of more instruction and experience. Also, it seems to pronunciation teachers that students make considerable progress in producing English vowel during the first year of their studies. The question is why this progress is not paralleled in perception. Do students have relatively good perception at the beginning of their studies (as the results showed) and is the only progress they make in faithful production? It is as if they had the perceptual sensitivity and skills ready at the onset of their first year at university as a result of previous experience and that pronunciation classes only gave them the articulatory skills to produce the vowels in a more native-like manner. Production studies among Polish advanced learners of English and university students of English have been conducted (Gonet et al. (2010a,b), Lipińska 2017, Porzuczek 2010, Rojczyk (2010a,b, 2011), Schwartz et al. (2016) and Szpyra-Kozłowska et al. (2014)). To determine whether this level of perception is enough to benefit from articulatory instruction and that is why perception does not significantly progress, or whether in the case of intensive articulatory training in a formal setting, production can, at least for some phones, precede perception, we would need parallel perception and production longitudinal studies.

Special attention should be paid to understanding the role of duration in non-native speech. Previous research has provided evidence for Bohn's (1995) desensitization hypothesis (for example, Bohn 1995, Flege et al. 1997 and Cebrian 2006), the feature hypothesis (McAllister et al. 2002) and actually for rejecting both of them (Lengeris 2009). The results of study two show complex and generally little reliance on duration: no influence on discrimination or (dis-)similarity ratings, lower goodness ratings for long/free vowels (but then spectral differences accompanied duration differences), as well as the influence of duration on English vowel identification. Some effects of paying closer attention to the duration of the vowel stimuli by learners of English and Dutch, but not by learners of French, were evident in study three. These results indirectly support the feature hypothesis (McAllister et al. 2002), as native Polish listeners did not pay attention to duration unless encouraged by a task or sensitized to duration in the L2 English or L3 Dutch, as shown in the Dutch vowel assimilations to vowel plus glide sequences in study three. The length feature is not used contrastively in Polish, so Polish listeners find it difficult in non-native vowel perception. Lengeris (2009) concludes his study by claiming that L2 vowel perception relies on an interplay between spectral and temporal cues, but temporal cues do not have a special status. The current study shows that duration has an important role in some tasks, that sensitivity to duration can be boosted by the L2 and L3 and surface in the L3 or foreign vowel perception, and that listeners pay more attention to duration than to tongue advancement cues. The non-obvious role of duration should be further studied to find the contexts in which it starts to play a role, and to determine under what circumstances the attention to duration is transferred from the L2 to the L3 or foreign language vowel perception.

Also, in order to encompass the development of speech perception by learners in a formal instruction setting, we would need to compare it with the development of perception in a naturalistic immersion setting for learners matched for age and language proficiency.

Defining a feature seems to be less problematic than defining a category. In the three studies presented in this book, sensitivity to features was continuous and varied from task to task (i.e. duration differences did not play a role in discrimination and (dis-)similarity ratings in study one, but they mediated identifications and misidentifications in study two). Features do not seem to be functioning as magnets, which subsume all the similar elements, whereas categories do act like magnets. Studies employing manipulated synthetic stimuli would be needed to investigate the role of each feature and determine whether or when a given feature assumes a categorical function. Studies of neural processing during non-native speech perception could verify current assumptions on category formation. Functional neuroimaging techniques can help us determine the details of the mechanism of representation and supplement the results of behavioral studies (Abutalebi and Green 2007; Bialystok, Craik, Grady, Chau, Ishii and Gunji 2005).

Another puzzling issue is the perception of front rounded vowels by speakers of languages which lack them and the perception of back unrounded vowels by speakers of languages which have or do not have front rounded vowels.

The studies could be also be replicated with accompanying visual inputs either matching or not matching the lip rounding in the auditory stimulus. According to the McGurk effect (McGurk and MacDonald 1976), the visual input overrides the auditory input. If subjects saw on the display a person producing a vowel without lip rounding and heard a high front rounded vowel, would they be more likely to choose a high front unrounded vowel and would they match the height of the chosen stimulus?

6.4. Implications for teaching pronunciation

The presented studies were meant to allow for more precision in explaining and predicting vowel assimilations in non-native languages. The existing models, enriched by the considerations related to phonetic features and their combinations, allow for a more precise explanation of non-native speech perception. The interpretation of the results obtained in the three studies in this book can inform foreign language learning and teaching specialists by drawing their attention to English contrasts which are perceived in a non-native like manner, which can in turn lead to misunderstandings in communication. Also, teaching the pronunciation of languages with front rounded vowels can make use of the present results.

Pronunciation teachers can assume that all vowel contrasts which have height differences should be relatively easy for learners to perceive. Duration does not have as privileged a status as Bohn's (1995) hypothesis seems to suggest, and learning to perceive duration differences in the L2 for learners with an L1 which does not use duration contrastively is a complex process, which takes place in stages. Any feature combinations which are more marked than in the L1 or other known languages are relatively challenging.

For Polish advanced learners of English, the perception of /iː-ɪ/ does not seem to be as big a problem, as their production could suggest. They discriminate excellently between the two sounds. Categorization and identification of /ɪ/ is only problematic in the velar context, which stems from phonotactic restrictions in Polish. Production problems in the case of this contrast are not related to perception in the environments not influenced by native phonotactics. The major cause of pronunciation problems probably lies in misleading orthography.

Clearly, English low vowels are challenging for Polish learners, whose L1 has one low vowel only. All the three English /æ, ʌ/ and /ɑː/ vowels are assimilated to Polish /a/, they are poorly discriminated from one another, judged to be similar and often confused. They differ in tongue advancement, which is the most difficult feature to perceive. Their perception has nevertheless improved throughout the year: goodness ratings became significantly lower, and /ɑː - ʌ/ and /ʌ - ɒ/ were judged to be less similar in stage two in comparison to stage one. These findings suggest that progress is possible even in the case of Single Category assimilations differentiated by tongue advancement. Nevertheless, the results of study one also clearly show that the English low vowels were perceived with difficulty even after two semesters of pronunciation training, so they need more attention.

The high central slightly rounded /uː/ and mid-high central slightly rounded /ʊ/ were also assimilated as Single Category, they were much better identified than the low vowels (the problems were related to the allophonic variation – English central vowels are fronted before alveolars), they were poorly discriminated and judged to be similar, but in the last two tests significant improvement in perception was observed. We can speculate that the high identification scores, especially in comparison with the low vowels, stem from the fact that there are only two English vowels assimilated to the Polish /u/ and that the difference between them lies both in height and duration. Similarly, as in the case of low vowels – more attention should be paid to the /uː - ʊ/ contrast.

The research in this book has not focused on teaching pronunciation, but in addition to traditional articulatory practice, raising metacompetence and awareness (Venkatagiri and Levis 2007, Kennedy and Trofimovich 2010, Wrembel 2005 and 2011), bimodal or enhanced bimodal training has recently been recommended specifically for perception practice (Escudero and Williams 2014). Distributional learning is learning based on a statistical learning mechanism, in which the relative frequency of exposure to tokens from the ends of a two category vowel continuum, or even their exaggerated values, improves discrimination of the two non-native contrasts. This method seems to be a reasonable choice for enhancing sensitivity to English low vowels and high and high-mid central vowels. Escudero, Benders and Lipsky (2009) also show how learner profiling helps determine the cue weighting typical for a given learner and how diagnosing students as sequential or interactive learners can help adjust the level of meaning-driven learning (van Leussen and Escudero 2015). These methods seem to be potentially very useful for advanced Polish learners of English. The only prerequisite for using them would be designing appropriate software or mobile applications.

The experiments in studies two and three showed that perceiving marked combinations of features is generally difficult, and it seems that by default Polish listeners associate rounding with backness. In pronunciation training in languages with front rounded vowels, this association needs to be reinterpreted so that front vowel contrasts can be formed on the basis of lip rounding.

6.5. Final remarks

At least some aspects of non-native vowel perception have been widely researched, while some, such as third language vowel perception or foreign language vowel perception, are still scarcely explored. This contribution has aimed at examining the idea that in addition to categories acting as magnets, it is selective attention to features that contributes to the process of non-native speech perception. In order to draw conclusions on the extent of selective attention to features, we tried to isolate its effects on vowel perception. This was not a straightforward task, if we take into account that all the vowels in the tested languages have their own sets or bundles of characteristics and separation of the effects of one feature from the other is not always viable. The role of selective attention to features has, not surprisingly, turned out to be of secondary importance, but we have observed its traces in the results of all the three studies. The features interacting with categorical perception were both universal (e.g. height) and experience-based. Since advanced learners were tested, we have also seen that learning to employ some of the features (height) is easier than others (backness distinctions). We have seen that perception development is a very gradual process. Language-specific bilingual sensitivity to features constitutes the major argument for selective attention to features in non-native vowel perception. If categories act like magnets, then features act in much more subtle yet fairly systematic ways.

Appendix. Mean per cent assimilation of Turkish and Dutch vowels to Polish categories. Response categories: Polish vowels and vowel plus glide combinations

Stimulus	L2	i /i/	ij /ij/	ił /iw/	y /ɨ/	e /e/	ej /ej/	eł /ew/	a /a/	aj /aj/	ał /aw/	o /ɔ/	u /u/	uj /uj/	uł /uw/
ɯ	English	–	–	–	27.84	7.45	–	–	20.00	–	–	20.39	24.31	–	–
		–	–	–	3.35	2.11	–	–	2.88	–	–	4.63	3.4	–	–
	French	–	0.32	–	29.84	15.56	–	–	15.56	0.32	–	34.60	3.81	–	–
		–	4	–	3.06	2.9	–	–	3.29	2	–	5.45	2.92	–	–
	Dutch	1.03	–	–	54.36	26.15	–	–	11.54	–	–	2.05	4.87	–	–
		3.5	–	–	3.01	4.34	–	–	3.4	–	–	3.75	4.58	–	–
œ	English	2.35	0.39	–	21.57	32.16	0.78	–	12.16	–	–	13.73	16.86	–	–
		1.33	3	–	3.2	3.51	3.5	–	2.97	–	–	4.31	3.42	–	–
	French	0.32	–	–	16.19	43.49	–	0.32	11.11	–	0.32	24.76	3.49	–	–
		1	–	–	2.82	3.67	–	7	3.4	–	5	5.44	3.18	–	–
	Dutch	0.77	–	–	45.90	43.08	0.51	0.77	3.59	–	–	0.00	5.13	–	0.26
		3.67	–	–	3.13	4.76	3.5	2.33	2.71	–	–	–	4.75	–	2

Stimulus	L2	i /i/	ij /ij/	ił /iw/	y /ɨ/	e /e/	ej /ej/	eł /ew/	a /a/	aj /aj/	ał /aw/	o /ɔ/	u /u/	uj /uj/	ł /uw/
ɑ	English	–	–	–	0.39	–	–	–	62.75	–	0.39	35.69	0.78	–	–
		–	–	–	3	–	–	–	4.13	–	7	5.9	3	–	–
	French	–	–	–	–	–	–	–	46.98	–	0.32	52.38	0.32	–	–
		–	–	–	–	–	–	–	5.78	–	2	5.71	5	–	–
	Dutch	–	–	–	–	–	–	–	88.97	–	–	11.03	–	–	–
		–	–	–	–	–	–	–	5.73	–	–	5.98	–	–	–
aː	English	–	–	–	–	0.39	–	–	96.86	0.39	0.39	1.96	–	–	–
		–	–	–	–	2	–	–	5.59	5	4	4.6	–	–	–
	French	–	–	–	–	–	–	–	98.10	0.32	0.95	0.63	–	–	–
		–	–	–	–	–	–	–	5.64	7	3.33	2.5	–	–	–
	Dutch	–	–	–	–	–	–	–	99.49	–	0.51	–	–	–	–
		–	–	–	–	–	–	–	5.08	–	2	–	–	–	–
ɛ	English	–	–	–	–	96.47	–	–	3.53	–	–	–	–	–	–
		–	–	–	–	5.17	–	–	4.33	–	–	–	–	–	–
	French	–	–	–	1.27	71.11	–	–	26.03	–	0.63	0.32	0.32	0.32	–
		–	–	–	4	5.32	–	–	4.90	–	3.5	4	5	7	–
	Dutch	–	–	–	–	99.23	–	0.26	0.51	–	–	–	–	–	–
		–	–	–	–	5.66	–	1	2.5	–	–	–	–	–	–

Stimulus	L2	i /i/	ij /ij/	ił /iw/	y /ɨ/	e /e/	ej /ej/	eł /ew/	a /a/	aj /aj/	ał /aw/	o /ɔ/	u /u/	uj /uj/	uł /uw/
eː	English	3.92	2.35	1.96	5.10	52.16	32.55	–	0.78	–	–	0.39	–	–	0.78
		4	5	4.8	3.46	4.80	4.73	–	6	–	–	1	–	–	3.5
	French	5.40	0.32	0.32	4.44	68.57	20.32	0.32	0.32	–	–	–	–	–	–
		2.65	2	4	6.29	4.89	5.53	3	1	–	–	–	–	–	–
	Dutch	7.44	3.85	–	0.26	21.54	66.41	0.26	–	0.26	–	–	–	–	–
		5.90	2.47	–	3	2.99	4.13	4	–	1	–	–	–	–	–
yː	English	1.18	3.14	0.00	5.88	0.39	0.00	0.00	0.00	0.00	0.00	0.39	62.35	14.90	11.76
		2.67	3	–	3.6	1	–	–	–	–	–	7	4.12	3.82	4.4
	French	0.63	0.00	0.32	2.22	0.00	0.00	0.32	0.00	0.00	0.00	0.00	84.76	8.57	3.17
		2.5	–	6	2.29	–	–	7	–	–	–	–	4.20	4.56	5.4
	Dutch	4.87	0.00	0.26	12.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	59.49	5.38	17.95
		2	–	3	3.26	–	–	–	–	–	–	–	2.89	3.14	3.24
ʏ	English	–	–	–	25.10	0.78	0.39	–	0.39	–	–	–	70.59	0.39	2.35
		–	–	–	4.05	3	4	–	4	–	–	–	4.34	4	4
	French	0.32	–	–	36.83	4.13	–	–	0.95	–	–	1.27	53.65	0.32	2.54
		1	–	–	3.31	2.69	–	–	2.33	–	–	4	4.27	5	3.63
	Dutch	–	–	–	55.90	–	–	–	0.26	–	–	–	43.59	–	0.26
		–	–	–	3.65	–	–	–	1	–	–	–	3.8	–	2

Stimulus	L2	i /i/	ij /ij/	ił /iw/	y /ɨ/	e /e/	ej /ej/	eł /ew/	a /a/	aj /aj/	ał /aw/	o /ɔ/	u /u/	uj /uj/	uł /uw/
øː	English	0.39	–	–	18.43	4.71	–	2.75	0.39	–	–	0.39	55.29	3.92	13.73
		3	–	–	3.26	2.92	–	2.86	5	–	–	3	4.43	4.1	4.49
	French	–	–	0.32	31.43	10.79	–	2.22	0.32	–	0.63	1.59	43.49	1.90	7.30
		–	–	1	3.53	3.53	–	4.29	4	–	5	2	3.85	4.17	4.61
	Dutch	–	–	–	40.77	0.51	–	0.77	0.26	–	0.26	1.28	41.79	1.03	13.33
		–	–	–	3.11	1.5	–	1.33	2	–	2	3.4	3.19	3.75	4.06
uː	English	–	–	–	–	–	–	–	–	–	–	–	90.20	1.57	8.24
		–	–	–	–	–	–	–	–	–	–	–	4.913043478	4.25	4.857142857
	French	–	–	0.32	–	0.32	–	–	0.32	–	–	0.32	91.43	–	7.30
		–	–	2	–	1	–	–	4	–	–	7	5.20	–	6.09
	Dutch	–	–	–	1.79	–	–	–	–	–	–	–	85.38	–	12.82
		–	–	–	2.86	–	–	–	–	–	–	–	4.4	–	3.98

References

Abutalebi, Jubin and David Green. 2007. Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics* 20(3): 242-275. DOI: 10.1016/j.jneuroling.2006.10.003

Adank, Patti, Roeland van Hout and Roel Smits. 2004. An acoustic description of the vowels of Northern and Southern Standard Dutch. *Journal of the Acoustical Society of America* 116(3): 1729-1738. DOI: 10.1121/1.1779271

Alispahic, Samra, Karen E. Mulak and Paola Escudero. 2017. Acoustic properties predict perception of unfamiliar Dutch vowels by adult Australian English and Peruvian Spanish listeners. *Frontiers in Psychology* 8:52. DOI: 10.3389/fpsyg.2017.00052

Antoniou, Mark, Michael D. Tyler, Catherine T. Best. 2012. Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics* 40: 582-594. DOI: 10.1016/j.wocn.2012.05.005

Antoniou, Mark, Catherine T. Best, Michael D. Tyler. 2013. Focusing the lens of language experience: Perception of Ma'di stops by Greek and English bilinguals and monolinguals. *Journal of the Acoustical Society of America* 133(4): 2397-2411. DOI: 10.1121/1.4792358

Antoniou, Mark, Eric Liang, Marc Ettlinger and Patrick C. M. Wong. 2015. The bilingual advantage in phonetic learning. *Bilingualism: Language and Cognition* 18(4): 683-695. DOI: 10.1017/S1366728914000777

Aoki, Rika and Fumiaki Nishihara. 2013. Sound feature interference between two second languages: An expansion of the feature hypothesis to the multilingual situation in SLA. In *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society*, ed. Chundra Cathcart, I-Hsuan Chen, Greg Finley, Shinae Kang, Clare S. Sandy and Elise Stickles, 18-31. https://journals.linguisticsociety.org/proceedings/index.php/-BLS/article/viewFile/847/630

Backley, Philip. 2011. *An introduction to Element Theory.* Edinburgh: Edinburgh University Press.

Balas, Anna. 2009. Why can Poles perceive Sprite but not Coca-Cola? A Natural Phonological account. In *Phonology in Perception,* ed. Paul Boersma and Silke Hamann, 25-53. Berlin: Mouton de Gruyter.

Balas, Anna. 2017a. The influence of second language vowels on foreign language vowel perception. *Linguistic Society of America Proceedings* 2, ed. Patrick Farrell, 44: 1-5. DOI: 10.3765/plsa.v2i0.4085

Balas, Anna. 2017b. More evidence for selective attention to features: Perception of the Turkish high back unrounded vowel. In *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017,* ed. Elena Babatsouli, 65-70. http://ismbs.eu/publications-2017

Balas, Anna. 2018. English vowel perception by Polish advanced learners of English. *Canadian Journal of Linguistics* 63 (3): 1-30. DOI: 10.1017/cnj.2018.5

Bałutowa, Bronisława. 1965. *Wymowa angielska dla wszystkich*. Warszawa: Wiedza Powszechna.

Bardel, Camilla and Ylva Falk. 2007. The role of the L2 in L3 acquisition: The case of Germanic syntax. *Second Language Research* 23(4): 459-484. DOI: 10.1177/0267658307080557

Bardel, Camilla and Ylva Falk. 2012. The L2 status factor and the declarative/procedural distinction. In *Third Language Acquisition in Adulthood,* ed. Jennifer Cabrelli Amaro, Suzanne Flynn and Jason Rothman, 61-78. Philadelphia: John Benjamins.

Bartolotti, James and Viorica Marian. 2012. Language learning and control in monolinguals and bilinguals. *Cognitive Science* 36(6): 1129-1147. DOI: 10.1111/j.1551-6709.2012.01243.x

Best, Cathrine. T. 1993. Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In *Developmental neurocognition: Speech and face processing in the first year of life*, ed. Benedicte De Boysson-Bardies, Scania de Schonen, Peter Jusczyk, Peter Mac-Neilage and John Morton, 289-304. Dordrecht, Netherlands: Kluwer Academic Publishers.

Best, Catherine T. 1994a. Learning to perceive the sound pattern of English. In *Advances in Infancy Research* (Vol. 9.), ed. Carolyn Rovee-Collier and Lewis P. Lipsitt, 217-304. Norwood, NJ: Ablex.

Best, Catherine. T. 1994b. The emergence of native-language phonological influences in infants: A Perceptual Assimilation Model. In *The development of speech perception: The transition from speech sounds to spoken words*, ed. Judith C. Goodman and Howard C. Nusbaum, 167- 244. Cambridge MA: MIT Press.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. Winifred Strange,171-204. Baltimore: York Press.

Best, Catherine T., Gerald W. McRoberts and Elizabeth Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America* 109(2): 775-794.

Best, Catherine T., Gerald W. McRoberts and Nomathemba M. Sithole. 1988. Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Perception and Performance* 14(3): 345-360. DOI: 10.1037/0096-1523.14.3.345

Best, Catherine T, Pierre Halle, Ocke-Schwen Bohn and Alice Faber. 2003. Cross-language perception of nonnative vowels: Phonological and phonetic effects of listeners' native languages. In *Proceedings of the 15th International Congress of Phonetic Sciences, Barcelona, Spain, August 3-9, 2003*, ed. M. J. Solé, D. Recasens, and J. Romero. Barcelona. http://www.internationalphoneticassociation.org/icphs/icphs2003

Best, Catherine T. and Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20: 305-330.

Best, Catherine T and Michael D. Tyler. 2007. Nonnative and second-language speech perception: Commonalities and complementarities. In *Second language speech learning* ed. Murray Munro and Ocke-Schwen Bohn, 13–34. Amsterdam: John Benjamins Publishing.

Bialystok, Ellen, Fergus I.M. Craik, Cheryl Grady, Wilkin Chau, Ryouhei Ishii, Atsuko Gunji and Christo Pantev. 2005. Effect of bilingualism on cognitive control in the Simon task: Evidence from MEG. *NeuroImage* 24(1): 40-49. DOI: 10.1016/j.neuroimage.2004.09.044

Bjelaković, Andrej. 2017. The vowels of contemporary RP: Vowel formant measurements for BBC newsreaders. *English Language and Linguistics* 21(3): 501-532. DOI:10.1017/S1360674316000253

Boersma, Paul and David Weenink. 2015. Praat: doing phonetics by computer [Computer program]. Retrieved from http://www.praat.org/.

Bogacka [Balas], Anna. 2004. On the perception of English high vowels by Polish learners of English. In *CamLing 2004: Proceedings of the University of Cambridge Second Postgraduate Conference in Language Research*, ed. Evangelia Daskalaki, Napoleon Katsos, Marios Mavrogiorgos and Matthew Reeve, 43-50. Cambridge: University of Cambridge.

Bogacka [Balas], Anna, Geoffrey Schwartz, Paulina Zydorowicz, Monika Połczyńska-Fiszer and Paula Orzechowska. 2006. The production and perception of schwa in second language acquisition: The case of Polish learners of English. In *IFAtuation: A life in IFA. A Festschrift for Professor Jacek Fisiak on the occasion of his 70th birthday*, ed. Katarzyna Dziubalska-Kołaczyk, 71-84. Poznań: Wydawnictwo Naukowe UAM.

Bongaerts, Theo. 1999. Ultimate attainment in L2 pronunciation: The case of very advanced late L2 learners. In *Second language acquisition and the critical period hypothesis*, ed. David Birdsong. Mahway, N.J.: Lawrence Erlbaum. 133-159.

Bohn, Ocke-Schwen. 1995. Cross-language speech perception in adults: First language transfer doesn't tell it all. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. Winifred Strange, 279-304. Timonium, MD: York Press.

Bohn, Ocke-Schwen and Catherine T. Best. 2012. Native language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *Journal of Phonetics* 40: 109-128. DOI: 10.1016/j.wocn.2011.-08.002

Bohn, Ocke-Schwen and James Emil Flege. 1990. Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics* 11: 303-328. DOI: 10.1017/S0142716400008912

Bohn, Ocke-Schwen and James Emil Flege. 1992. The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition* 14: 131-158. DOI: 10.1017/S0272263100010792

Bohn, Ocke-Schwen and James Emil Flege. 1997. Perception and production of a new vowel category by adult second language learners. In *Second Language Speech, Structure and Processes* ed. Allan James and Jonathan Leather, 11-52. Berlin: Mouton de Gruyter.

Brière, Eugène. 1966. An investigation of phonological interference. *Language* 42: 768-796. DOI: 10.2307/411832

Bundgaard-Nielsen, Rikke, Catherine T. Best, Michael D. Tyler. 2011. Vocabulary size matters: the assimilation of second-language Australian English vowels to first-language Japanese vowel categories. Applied Psycholinguistics 32: 51-67. DOI: 10.1017/S0142716410000287

Bundgaard-Nielsen, Rikke, Alexander Kilpatrick and Brett Baker. 2016. Japanese phonotactics influence perception of English consonants by Japanese learners of English. Paper presented at New Sounds, 8th International Symposium on the Acquisition of Second Language Speech, Aarhus, Denmark.

Cabrelli Amaro, Jennifer. 2016. Testing the Phonological Permeability Hypothesis: L3 phonological effects on L1 versus L2 systems. *International Journal of Bilingualism.* Advance Online Publication. DOI: 10.1177/1367006916637287.

Cabrelli Amaro, Jennifer and Jason Rothman. 2010. L3 acquisition and phonological permeability: A new test case for debates on the mental representation of non-native phonological systems. *International Review of Applied Linguistics in Teaching, IRAL* 48(2/3): 273-294.

Cawalho, Ana Maria and Antonio Jose Bacelar da Silva. 2006. Cross-linguistic influence in third language acquisition: The case of Spanish-English bilinguals' acquisition of Portuguese. *Foreign Language Annals*, 39(2): 185-202. DOI: 10.1111/j.1944-9720.2006.tb02261.x.

Cebrian, Juli. 2002. Phonetic similarity, syllabification and phonotactic constraints in the acquisition of a second language contrast. Ph.D. dissertation. Toronto Working Papers in Linguistics Dissertation Series. Toronto, Canada: Department of Linguistics, University of Toronto.

Cebrian, Juli. 2006. Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics* 34(3): 372-387. DOI: 10.1016/j.wocn.2005.08.003

Cenoz, Jasone. 2001. The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* ed. Jasone Cenoz, Britta Hufsein and Ulrike Jessner, 8-20. Clevedon: Multilingual Matters.

Chen, Matthew. 1970. Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22: 129-159. DOI: 10.1159/000259312

Chládková, Kateřina and Václav Jonáš Podlipský. 2011. Native dialect matters: Perceptual Assimilation of Dutch vowels by Czech listeners. *Journal of the Acoustical Society of America* 130(4): EL 186-92. DOI: 10.1121/1.3629135

Clements, George Nick. 2003. Feature economy in sound systems. *Phonology*, 20, 287–333. DOI: 10.1017/S095267570400003X

Cohen, Stephen P., G. Richard Tucker and Wallace E. Lambert. 1967. The comparative skills of monolinguals and bilinguals in perceiving phoneme sequences. *Language and Speech* 10 (3): 159–168. DOI: 10.1177/002383096701000302
Collins, Beverly and Inger Mees. 1981. *The sounds of English and Dutch*. The Hague: Leiden University Press.
Collins, Beverly and Inger Mees. 2013. *Practical phonetics and phonology.* 3rd edition. London: Routledge.
Common European Framwork of Reference for Languages: Learning, Teaching, Assessment. 2011. Council of Europe, Education and Languages. (https://www.coe.int-/EN/web/common-european-framework-reference-languages/) (date of access: 15 Apr. 2016)
Crothers, John. 1978. Typology and universals of vowel systems. In *Universals of human language. Volume 2*, ed. Joseph Greenberg, Charles Ferguson and Edith Movarvcsik. Stanford, California: Stanford University Press.
Cruttenden, Alan. 2014. *Gimson's pronunciation of English*. 8th edition. New York: Routledge.
Crystal, Thomas H. and Arthur S. House. 1988. The duration of American-English vowels: An overview. *Journal of Phonetics* 16: 263-284.
Daunmu, San. 2016. *A theory of phonological features*. Oxford: Oxford University Press.
Davidiak, Elena. 2010. One? ¿Dos? Drei! A study of code-switching in child trilingualism. PhD thesis, University of Iowa. http://ir.uiowa.edu/etd/484.
De Angelis, Gessica. 2007. *Third or Additional Language Acquisition.* Clevedon: Multilingual Matters.
Deterding, David. 1997. The formants of monophthong vowels in standard southern British English pronunciation. *Journal of the International Phonetic Association* 27: 47–55. DOI: 10.1017/S0025100300005417
Dłuska, Maria. 1981. *Fonetyka polska* [Polish phonetics]. Warszawa [Warsaw]: Państwowe Wydawnictwo Naukowe.
Donegan, Patricia. 2001. Constraints and processes in phonological perception. In *Constraints and Preferences* ed. Katarzyna Dziubalska-Kołaczyk, 42-68. Berlin: Mouton de Gruyter.
Donegan, Patricia. 2015. The emergence of phonological representation. In *The Handbook of Language Emergence* ed. Brian MacWhinney and William O'Grady, 35-52. Chichester: John Willey & Sons, Inc. DOI: 10.1002/9781118346136
Dupoux Emmanuel and Sharon Pepperkamp. 2002. A typological study of stress 'deafness'. In *Laboratory Phonology* 7, ed. Carlos Gussenhoven and Natasha Warner, 203-240. Berlin: Mouton de Gruyter.
Dziubalska-Kołaczyk, Katarzyna. 1990. A natural model of acquisition of second language phonology and the notion of relative markedness. In *Naturalists at Krems*, ed. Julián Méndez Dosuna and Carmen Pensado, 43-51. Salamanca: Universidad de Salamanca.
Dziubalska-Kołaczyk, Katarzyna. 2016. Accents in the phonological context: In search of a big picture. (Plenary paper presented at Accents, Łódź, 1-3 December 2016.).

Dziubalska-Kołaczyk, Katarzyna and Magdalena Wrembel. 2017. Modeling the acquisition of foreign language speech: Old meets new. (Paper presented at the 50th Annual Meetign of the Societas Linguistica Europaea, 10-13 September 2017.).

Eckman, Fred. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning* 27: 315-330. DOI: 10.1111/j.1467-1770.1977.tb00124.x

Eckman, Fred. 1981. On predicting phonological difficulty in second language acquisition. *Studies in Second Language Acquisition* 4: 18-30. DOI: 10.1017/S0272263100-004253

Ellis, Rod. 1984. *Classroom second language development*. Oxford: Pergamon Press.

Elvin, Jaydene, Paola Escudero and Polina Vasiliev. 2014. Spanish is better than English for discriminating Portuguese vowels: Acoustic similarity versus vowel inventory size. *Frontiers in Psychology* 5: 1188. DOI: 10.3389/fpsyg.2014.01188

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

Escudero, Paola. 2006. “Second language phonology: The role of perception.” In: *Phonology in context,* ed. M. Pennington. New York, NY: Palgrave Macmillan. 109-134.

Escudero Paola. 2009. “The linguistic perception of similar L2 sounds.” In: *Phonolgoy in Perception,* eds. Paul Boersma and Silke Hamann. Berlin: Walter de Gruyter. 152-190.

Escudero, Paola, Titia Benders and Silvia C. Lipsky. 2009. Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics* 37(4): 452-465. DOI: 10.1016/j.wocn.2009.07.006

Escudero, Paola and Paul Boersma. 2004. Bridging the gap between L2 speech perception and phonological theory. *Studies in Second Language Acquisition* 26: 551-585. DOI: 10.1017/S0272263104040021

Escudero, Paola, Ellen Simon and 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. DOI: 10.1016/j.wocn.2011.11.004

Escudero Paola and Daniel Williams. 2012. Native dialect influences second-language vowel perception: Peruvian vs. Iberian Spanish learners of Dutch. *Journal of the Acoustical Society of America* 131(5): EL406-EL412. DOI: 10.1121/1.3701708.

Escudero, Paola and Daniel Williams. 2014. Distributional learning has immediate and long-lasting effects. *Cognition* 133(2): 408-413. DOI: 10.1016/j.cognition.2014.07.002

Fabricius, Anne H. 2007. Variation and change in the the TRAP and STRUT vowels of RP: a real time comparison of five acoustic data sets. *Journal of the International Phonetic Association* 37(3): 293-320. DOI: 10.1017/S002510030700312X

Falk, Ylva and Camilla Bardel. 2010. The study of the role of the background languages in third language acquisition: The state of the art. *International Review of Applied Linguistics in Language Teaching* 48: 185-220. DOI: 10.1515/iral.2010.009

Falk, Ylva and Camilla Bardel. 2011. Object pronouns in German L3 syntax: Evidence for the L2 status factor. *Second Language Research* 27(1): 59-82. DOI: 10.1177/0267658310386647

Faris, Mona M, Catherine T. Best and Michael Douglas Tyler. 2016. An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *The Journal of the Acoustical Society of America* 139(1): EL1-EL5. DOI: 10.1121/1.4939608

Fernandes-Boëchat, Marcia Helena. 2007. The CCR Theory: A cognitive strategy research proposal for individual multilingualism. *Revista Luminária* 8(1).

Flege, James Emil. 1987a. Effects of equivalence classification on the production of foreign language speech sounds. In *Sounds Patterns in Second Language Acquisition* ed. Allan James and Jonathan Leather, 9-39. Dordrecht: Foris Publications.

Flege, James Emil. 1987b. The production of new and similar phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15, 47-65. DOI: 10.1017/S0272263100010792

Flege, James Emil. 1991. Orthographic evidence for the perceptual identification of vowels in Spanish and English. *The Quarterly Journal of Experimental Psychology* 43A: 701-731. DOI: 10.1080/14640749108400993

Flege, James Emil. 1995. Second language speech learning: Theory, problems, findings. In *Speech perception and linguistic experience: Issues in cross- language research* ed. Winifred Strange. 233-277. Timonium, MD: York Press.

Flege, James Emil and Wieke Eefting. 1987. Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech and Communication* 6: 185-202.

Flege, James Emil, Ocke-Schwen Bohn and Sunyoung Jang. 1997. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25: 437-470. DOI: 10.1006/jpho.1997.0052

Flege, James Emil, Murray J. Munro and Robert Allen Fox. 1994. Auditory and categorial effects on cross-language vowel perception. *Journal of the Acoustical Society of America* 95: 3623-3641. DOI: 10.1121/1.409931

Flege, James Emil, Naoyuki Takagi and Virginia Mann. 1996. Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Journal of the Acoustical Society of America* 99(2): 1161-1173.

Flynn, Suzanne, Clair Foley and Inna Vinnitskaya. 2004. The Cumulative-Enhancement Model for language acquisition: Comparing adults' and children's patterns of development in L1, L2 and L3 acquisition of relative clauses. *The International Journal of Multilingualism* 1(1): 3-16. DOI: 10.1080/14790710408668175

Fougeron, Cécile and Caroline L. Smith. 1993. French. *Journal of the International Phonetic Association* 23(2): 73-76. 10.1017/S0025100300004874

Fox, Robert A., James E. Flege and Murray J. Munro. 1995. The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustical Society of America* 97(4): 2540-2551.

Fry, D.B. 1955. Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America* 27: 765-768. DOI: 10.1121/1.1908022

Fry, D.B. 1958. Experiments in the perception of stress. *Language and Speech* 1 (2): 126-152. DOI: 10.1177/002383095800100207

Gallardo del Puerto, Francisco. 2007. Is L3 phonological competence affected by the learner's level of bilingualism? *International Journal of Multilingualism* 4(1): 1-16. DOI: 10.2167/ijm042.0

Ghaffarvand Mokari, Payam and Stefan Werner. 2017. Perceptual assimilation predicts acquisition of foreign language sounds: The case of Azerbaijani learners' production and perception of Standard Southern English vowels. *Lingua* 185: 81-95. DOI: 10.1016/j.lingua.2016.07.008

Giegerich, Heinz J. 1992. *English Phonology: An Introduction*. Cambridge: Cambridge University Press.

Gimson, Alfred Charles. 1964. Phonetic change and the RP vowel system. In *In honour of Daniel Jones: Papers contributed on the occasion of his eightieth birthday*, ed. Daniel Jones, David Abercrombie, Dennis Butler Fry, Peter MacCarthy and Norman Carson Scott, 131-136. London: Longman.

Gimson, Alfred Charles. 1980. *An introduction to the pronunciation of English*. London: Edward Arnold.

Gimson, Alfred Charles. 1984. The RP accent. In *Language in the British Isles*, ed. Peter Trudgill, 45-54. Cambridge: Cambridge University Press.

Gonet, Wiktor, Jolanta Szpyra-Kozłowska and Radosław Święciński. 2010a. Clashes with ashes. In *Issues in Accensts of English 2: Variability and Norm,* ed. Ewa Waniek-Klimczak. 213-228. Newcastle upon Tyne: Cambridge Scholars Publishing.

Gonet, Wiktor, Jolanta Szpyra-Kozłowska and Radosław Święciński. 2010b. The acquisition of vowel reduction by Polish students of English. In *Issues in Accensts of English 2: Variability and Norm,* ed. Ewa Waniek-Klimczak. 291-307. Newcastle upon Tyne: Cambridge Scholars Publishing.

Goto, Hiromu. 1971. Auditory perception by normal Japanese adults of the sounds "l" and "r." *Neuropsychologia* 9: 317-323. DOI:10.1016/0028-3932(71)90027-3

Gottfried, Terry. 1984. Effects of consonant context on the perception of French vowels. *Journal of Phonetics* 12: 91-114.

Gut, Ulrike. 2010. Cross-linguistic influence in L3 phonological acquisition. *International Journal of Multilingualism* 7(1): 19-38. DOI: 10.1080/14790710902972248

Hammarberg, Björn. 2010. The languages of the multilingual: Some conceptual and terminological issues. *International Review of Applied Linguistics, IRAL* 48: 91-104.

Hammarberg, Björn and Britta Hammarberg. 1993. Articulatory re-setting in the acquisition of new languages. *Phonum* 2: 61-67.

Hammarberg, Björn and Britta Hammarberg. 2005. Re-setting the basis of articulation in the acquisition of new languages: A third-language case study. In *Introductory Readings in L3* ed. Britta Hufsein and Robert Fouser, 11-18. Tübingen: StauFFenburg Verlag.

Halle, Morris and Kenneth N. Stevens. 1969. On the feature “Advanced Tongue Root”. *MIT Research Laboratory of Electronics Quarterly Progress Report* 94: 209-215.

Hallé, Pierre A., Catherine T. Best and Andrea Levitt. 1999. Phonetic vs. phonological influences on French listeners’ perception of American English approximants. *Journal of Phonetics* 27(3): 281-306. DOI: 10.1006/jpho.1999.0097

Hannisdal, Bente. 2007. *Variability and change in Received Pronunciation: a study of six phonological variables in the speech of television newsreaders*. Ph.D. dissertation. Universitet i Bergen.

Hawkins, Sarah and Jonathan Midgley. 2005. Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association* 35: 183-199. DOI:10.1017/S0025100305002124

House, Arthur S. 1961. On vowel duration in English. *Journal of the Acoustical Society of America* 33: 1174. DOI: 1178.10.1121/1.1908941

House, Arthur S. and Grant Fairbanks. 1953. The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25: 105-113. DOI: 10.1121/1.1906982

Hughes, Arthur, Peter Trudgill, Dominic Watt. 2012. English accents and dialects: An introduction to social and regional varieties of English in the British Isles. 5th edition. London: Routledge.

Ingram, John CL and See-Gyoon Park. 1997. Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics* 25(3): 343-370. DOI: 10.1006/jpho.1997.0048

Iverson, Paul, Patricia Kuhl, Reiko Akahane-Yamada, Eugen Diesch, Yoh’ich Tohkura, Andreas Kettermann and Claudia Siebert. 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87(1): B47-B57.

Iverson, Paul and Bronwen G. Evans. 2007. Learning English vowels with different first-language systems: Perception of formant targets, formant movement and duration. *Journal of the Acoustical Society of America* 122(5): 2842-2854. DOI: 10.1121/1.2783198

Jassem, Wiktor. 1972. *Fonetyka języka angielskiego*. Warszawa: Państwowe Wydawnictwo Naukowe.

Jassem, Wiktor. 1987. *The phonology of modern English*. Warszawa: Państwowe Wydawnictwo Naukowe.

Jassem. Wiktor. 2003. Polish. *Journal of the International Phonetic Association* 33: 103-107. DOI: 10.1017/S0025100303001191

Johnson, Keith, Edward Flemmin, Richard Wright. 1993. The hyperspace effect: Phonetic targets are hyperarticulated. *Language* 69(3): 505-528. DOI: 10.2307/416697

Kennedy, Sara and Pavel Trofimovich. 2010. Language awareness and second language pronunciation: A classroom study. *Language Awareness* 19(3): 171-185. DOI: 10.1080/09658416.2010.486439

Kiliç, Mehmet Akif and Faith Öğüt. 2004. A high unrounded vowel in Turkish: Is it a central or a back vowel? *Speech Communication* 43: 143-154. DOI: 10.1016/-j.specom.2004.03.001

Kopečková, Romana. 2015. Differences in the perception of English vowel sounds by child L2 and L3 learners. In *Universal or Diverse Paths to English Phonology* ed. Ulrike Gut, Robert Fuchs and Eva-Maria Wunder, 71-89. Berlin: Mouton de Gruyter.

Kornfilt, Jaklin. 1997. *Turkish.* London: Routledge.

Krashen, Stephen. 1981. *Second language acquisition and second language learning*. Oxford: Pergamon Press.

Krzeszowski, Tomasz. 1968. *An outline of American English phonetics*. Warszawa: Państwowe Wydawnictwo Naukowe.

Kuhl, Patricia. 1991a. Human adults and human infants show a "percetual magnet effect" for the prototypes of speech categories, monkeys for not. *Perception and Psychophysics* 50(2): 93-107.

Kuhl, Patricia. 1991b. Perception, cognition, and the ontogenetic and phylogenetic emergence in human speech. In *Plasticity of development*, ed. Steven Brauth, Robert J. Dooling and William S. Hall, 73-106. Cambridge, MA: MIT Press.

Kuhl, Patricia. 1994. Learning and representation in speech and language. *Current Opinion in Neurobiology* 4(6): 812-822.

Kuhl, Patricia. 2000a. A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America* 97(22): 11850-11857.

Kuhl, Patricia. 2000b. Language, mind, and brain: Experience alters perception. In *The new cognitive neurosciences*, ed. Michael S. Gazzaniga, 99-115. Cambridge, MA: MIT Press.

Kuhl, Patricia, Barbara T. Conboy, Sharon Coffey-Corina, Denis Padden, Maritza Rivera-Gaxiola and Tobey Nelson. 2008. Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 363(1493): 979-1000.

Kurath, Hans and Raven Ioor McDavid. 1961. *The pronunciation of English in the Atlantic States*. Ann Arbor: University of Michigan Press.

Labov, William, Sharon Ash and Charles Boberg. 2008. *The atlas of North American English.* Berlin: Mouton de Gruyter.

LaCharité, Darlene and Carole Paradis. 2005. Category preservation and proximity versus phonetic approximation in loanword adaptation. *Linguistic Inquiry* 36: 223-258. DOI: 10.1162/0024389053710666

Ladefoged, Peter. 1993. *A Course in Phonetics.* Fort Worth: Harcourt Brace Jovanovich College Publishers.

Lado, Robert. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: The University of Michigan Press.

Ladefoged, Peter and Keith Johnson. 2015. *A course in phonetics*. Belmont, CA: Cengage Learning.

Lengeris, Angelos. 2009. Perceptual assimilation and L2 learning: Evidence from the perception of Southern British English vowels by native speakers of Greek and Japanese. Phonetica 66: 169-187. DOI:10.1159/000235659

Lenneberg, Eric. 1967. *Biological foundations of language*. New York: John Wiley and Sons.

van Leussen, Jan Willem and Paula Escudero. 2015. Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology* 6: 1000. DOI: 10.3389/fpsyg.2015.01000

Levy, Erika S. and Winifred Strange. 2008. Perception of French vowels by American English adults with and without French language experience. *Journal of Phonetics* 36: 141-157. DOI: 10.1016/j.wocn.2007.03.001

Lindblom, Björn E. 1963. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America* 35: 1773-1781. DOI: 10.1121/1.1918816

Lindsey, Geoff. 2012. The British English vowel system. http://englishspeechservices.com/blog/british-vowels/

Lipińska, Dorota. 2017. Production of English /iː/ and /ɪ/ by Polish learners of English: A longitudinal study. In *Proceedings of the International Symposium on Monolingual and Bilingual Speech 2017*, ed. Elena Babatsouli, 180-185. http://ismbs.eu/publications-2017.

Maddieson, Ian. 1984. *Patterns of Sounds.* Cambridge: Cambridge University Press.

Major, Roy C. 2001. Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology. Mahwah, New Jersey: Lawrence Erlbaum.

Mayr, Robert and Paola Escudero. 2010. Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism: Language and Cognition* 13(3): 279-297. DOI: 10.1017/S1366728909990022 279

McAllister, Robert, James E. Flege and Thorsten Piske. 2002. The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Jounral of Phonetics* 30: 229-258. DOI: 10.1006/jpho.2002.0174

McGurk, Harry and John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264: 746-748. DOI: 10.1038/264746a0

Michaels, David. 1974. Some replacements and phonological systems. *Linguistics* 126: 69-81.

Miyawaki, Kuniko, James J. Jenkins, Winifred Strange, Alvin M. Liberman, Robert Verbrugge and Osamu Fujimura. 1975. An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics* 18(5): 331-340.

Morrison, Geoffrey Stewart and Peter F. Assmann, ed. 2013. *Vowel inherent spectral change.* Berlin: Springer-Verlag.

Moulton, William G. 1990. Some vowel systems in American English. In *Studies in the Pronunciation of English: A commemorative Volume in Honour of A.C. Gimson*, ed. Susan Ramsaran, 119-135. London: Routledge.

Nearey, Terrance. 1989. Static, dynamic, and relational properties in vowel perception. *Jounral of the Acoustical Society of America* 85: 2088-2113. DOI: 10.1121/1.2024120

Nimz, Katharina. 2016. *Sound Perception and Production in a Foreign Language: Does Orthography Matter?* Potsdam: Universitätsverlag Potsdam.

Ong, Jia Hoong, Denis Burnham, Catherine J. Stevens and Paola Escudero. 2016. Naïve learners show cross-domain transfer after distributional learning: The case of lexical and musical pitch. *Frontiers in Psychology* 7: 1189. DOI: 10.3389/fpsyg.2016.01189

Onishi, Hiromi. 2016. The effects of L2 experience on L3 perception. *International Journal of Multilingualism* 13(4): 459-475. DOI: 10.1080/14790718.2016.1217604

Polka, Linda and Ocke-Schwen Bohn. 2003. Asymmetries in vowel perception. *Speech Communication* 41: 221-231. DOI: 10.1016/S0167-6393(02)00105-X

Polka, Linda and Ocke-Schwen Bohn. 2011. Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics* 39: 467-478. DOI: 10.1016/j.wocn.2010.08.007

Pajak, Bozena and Roger Levy. 2014. The role of abstraction in non-native speech perception. *Journal of Phonetics* 46: 147-160. DOI: 10.1016/j.wocn.2014.07.001

Park, Hanyong and Kenneth J. de Jong. 2008. Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics* 36(2008): 704-723. DOI: 10.1016/j.wocn.2008.06.002

Patihis, Lawrence, Janet S. Oh and Tayopa Mogilner. 2015. Phoneme discrimination of an unrelated language: Evidence for a narrow transfer, but not a broad-based bilingual advantage. *International Journal of Bilingualism* 19(1): 3-16. DOI: 10.1177/1367006913476768

Peterson, Gordon E. and Ilse Lehiste. 1960. Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America* 32: 693-703. DOI: 10.1121/1.1908183

Polka, Linda and Ocke-Schwen Bohn. 2003. Asymmetries in vowel perception. *Speech Communication* 41: 221-231. DOI: 10.1016/S0167-6393(02)00105-X

Polka, Linda and Ocke-Schwen Bohn. 2011. Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics* 39: 467-478. DOI:10.1016/j.wocn.2010.08.007

Port, Robert F. and Adam Leary. 2005. Against formal phonology. *Language* 81: 927-964. DOI: 10.1353/lan.2005.0195

Porzuczek, Andrzej. 1996. The perception of the vocalic system of English by advanced Polish learners. In *Foreign Language Acquisition Studies*, ed. Janusz Arabski, 36-43. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Porzuczek, Andrzej. 1999. The main reasons for vowel recognition failures in Polish learners of English. In *On language theory and practice: In honour of Janusz Arabski on the occasion of his 60th birthday. Vol. 2 Language Acquisition, Learning and Theory*, ed. Marta Wysocka, 93-100. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Porzuczek, Andrzej. 2007. English vowel quantity in Polish learner's speech perception and production. In *PASE Papers 2007. Studies in Language and Methodology of Teaching Foreign Languages*, ed. Janusz. Arabski, Danuta Gabryś-Barker and Andrzej Łyda, 96-105. Katowice: PARA.

Porzuczek, Andrzej. 2010. The weak forms of *to* in the pronunciation of Polish learners of English. In *Issues in Accents of English*, ed. Ewa Waniek-Klimczak, 309-323. Newcastle upon Tyne: Cambridge Scholars Publishing.

Porzuczek, Andrzej, Arkadiusz Rojczyk and Janusz Arabski. 2016. *Praktyczny kurs wymowy angielskiej dla Polaków*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Rallo Fabra, Lucrecia and Joaquín Romero. 2012. Native Catalan learners' perception and production of English vowels. *Journal of Phonetics* 40(3): 491-508. DOI: 10.1016/j.wocn.2012.01.001

Rakerd, Brad, Robert R. Verbrugge and Donald P. Shankweiler. 1984. Monitoring for vowels in isolation and in consonantal context. *Journal of the Acoustical Society of America*, 76: 27-31. DOI: 10.1121/1.391114

Rakerd, Brad and Robert R. Verbrugge. 1985. Linguistic and acoustic correlates of the perceptual structure found in an individual scaling studies of vowels. *Journal of the Acoustical Society of America* 77: 296-301. DOI: 10.1121/1.392393

Raphael, Lawrence J. 1972. Preceeding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America* 51: 1296-1303. DOI: 10.1121/1.1912974

Raphael, L. 1975. The physiological control of durational differences between vowels preceding voiced and voiceless consonants in English. *Journal of Phonetics,* 3: 25-33.

Raphael, Lawrence J., Fredericka Bell-Berti, Rene Collier and Thomas Baer. 1979. Tongue position in rounded and unrounded front vowel pairs. *Language and Speech* 22: 37-48. DOI: 10.1177/002383097902200103

Reszkiewicz, Alfred. 1981. *Correct your English pronunciation.* Warszawa: Państwowe Wydawnictwo Naukowe.

Ringbom, Håkan. 1987. *The Role of the Mother Tongue in Foreign Language Learning*. Clevedon: Multilingual Matters.

Rochet, Bernard L. 1995. Perception and production of second-language speech sounds by adults. In *Speech Perception and Linguistic Experience: Issues in cross-language research.* Timonium: York Press.

Rojczyk, Arkadiusz. 2010a. Forming new vowel categories in second language speech: The case of Polish learners' production of English /ɪ/ and /e/. *Research in Language* 8: 85-97.

Rojczyk, Arkadiusz. 2010b. Production and perception of vowel 'ash' by Polish learners of English. In *Proceedings from the Sixth International Symposium on the Acquisition of Second Language Speech* [CD ROM version], ed. Katarzyna Dziubalska-Kołaczyk, Magdalena Wrembel and Małgorzata Kul.

Rojczyk, Arkadiusz. 2011. Production and perception of vowel /æ/ by Polish learners of English. In *Achievements and Perspectives in SLA of Speech: New Sounds 2010,* ed. Magdalena Wrembel, Małgorzata Kul and Katarzyna Dziubalska-Kołaczyk, 239-250. Frankfurt: Peter Lang.

Rojczyk, Arkadiusz. 2013. Vowel quality and duration as a cue to word stress for non-native listeners: Polish listeners' perception of stress in English. In *Teaching and researching English accents in native and non-native speakers*, ed. Ewa Waniek-Klimczak, Linda R. Shockey, 59-71. Berlin: Springer Verlag.

Rothman, Jason. 2010. On the typological economy of syntactic transfer: Word order and relative clause attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics* 48: (2-3): 245-273. DOI: 10.1515/iral.2010.011

Rothman, Jason. 2011. L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research* 27 (1). 107-127. DOI: 10.1177/0267658310386439

Rothman, Jason. 2013. Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Initial stages of L3 romance and beyond. In *Romance Languages and Linguistic Theory 2011: Selected Papers for 'Going Romance' Utrecht 2011* ed. Sergio Baauw, Frank Drijkoningen, Luisa Meroni and Manuela Pinto, 217-248. Amsterdam: John Benjamins. DOI: 10.1075/rllt.5.11rot

Rothman, Jason. 2015. Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language and Cognition* 18(2): 179-190. DOI: 10.1017/S136672891300059X

Rothman, Jason and Becky Halloran. 2013. Formal linguistic approaches to L3/Ln acquisition: A focus on morphosyntactic transfer in adult multilingualism. *Annual Review of Applied Linguistics* 33: 51-67. DOI: 10.1017/S0267190513000032

Sawala, Krzysztof, Tomasz Szczegóła, Michał Jankowski and Jarosław Weckwerth. 2016. *Say it right – Multimedialny kurs wymowy i słownictwa angielskiego wersja 4.3.* Poznań: SuperMemo World – Oficyna Wydawnicza Atena.

Schimd, Stefan. 1997. The naturalness differential hypothesis: Cross-linguistic influence and universal preferences in interlanguage phonology and morphology. *Folia Linguistica* 31: 331-348. DOI: 10.1515/flin.1997.31.3-4.331

Schwartz, Geoffrey, Grzegorz Aperliński, Kamil Kaźmierski and Jarosław Weckwerth. 2016. Dynamic targets in the acquisition of L2 English vowels. *Research in Language* 14(2): 181-202. DOI: 10.1515/rela-2016-0011

Schwartz, Geoffrey, Grzegorz Aperliński, Mateusz Jekiel and Kamil Malarski. 2016. Spectral dynamics in L1 and L2 vowel perception. *Research in Language* 14(1): 61-77. DOI: https://doi.org/10.1515/rela-2016-0004

Schwartz, Geoffrey, Kamil Kaźmierski, Jarosław Weckwerth, Mateusz Jekiel and Kamil Malarski. Submitted. Vowel dynamic in the acquisition of L2 English – a cross-sectional and longitudinal study of L1 Polish learners. DOI: 10.13140/RG.2.2.24905.65122

Sebastián-Gallés, Núria and Salvador Soto-Faraco. 1999. On-line processing of native and non-native phonemic contrasts in early bilinguals. *Cognition* 72: 111-123.

Sharewood-Smith, Michael and Eric Kellerman. 1986. Crosslinguistic influence in second language acquisition: An introduction. In *Crosslinguistic influence in second language acquisition* ed. Eric Kellerman and Michael Sharewood-Smith, 1-9. Oxford: Pergamon Press.

Simon, Elen, Mathijs Debaene and Mieke Van Herreweghe. 2015. The effect of L1 regional variation on the perception and production of standard L1 and L2 vowels. *Folia Linguistica*, 49: 521-553. DOI: 10.1515/flin-2015-0018

Sobkowiak, Włodzimierz. 2004. *English phonetics for Poles: A resource books for learners and teachers*. Poznań: Wydawnictwo Poznańskie.

Stampe, David. 1969. The acquisition of phonetic representation. *Proceedings of the Chicago Linguistic Society* 5: 433-444

Stampe, David and Patricia Donegan. 2009. Hypotheses of Natural Phonology. *Poznań Studies in Contemporary Linguistics* 45(1): 1-31. DOI: 10.2478/v10010-009-0002-x

Strange, Winifred, Reiko Akahane-Yamada, Rieko Kubo, Sonja A. Trent and Kanae Nishi. 2001. Effects of consonantal context on perceptual assimilation of American English vowels by Japanese listeners. *Journal of the Acoustical Society of America* 109 (4): 1691-1704. DOI: 10.1121/1.1353594

Strange, Winifred and Valerie L. Shafer. 2007. Speech perception in second language learners: The re-education of selective perception. In *Phonology and second language acquisition* ed. Jette G. Hansen Edwards and Mary L. Zampini, 153-191. Amsterdam: John Benjamins Publishing Company.

Strycharczuk, Patrycja and Peter Jurgec. 2008. Prosodic influences on formant frequencies of Polish vowels. Paper presented at the 3rd Newcastle Postgraduate Conference in Theoretical and Applied Linguistics.

Szpyra-Kozłowska, Jolanta. 2016. Perception? Orthography? Phonology? Conflicting forces behind the adaptation of English /ɪ/ loanwords into Polish. *Poznań Studies in Contemporary Linguistics* 52(1): 119-147. DOI: https://doi.org/10.1515/psicl-2016-0001

Szpyra-Kozłowska, Jolanta and Włodzimierz Sobkowiak. 2011. *Workbook in English phonetics: For Polish students of English.* Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.

Szpyra-Kozłowska, Jolanta, Sławomir Stasiak and Radosław Święciński. 2014. On the teachability of English allophonic distinctions to intermediate Polish learners. In *Recent Developments in Applied Phonetics,* ed. Jolanta Szpyra-Kozłowska, Ewa Guz, Piotr Steinbrich and Radosław Święciński, 217-238. Lublin: Wydawnictwo KUL.

Trubetzkoy, Nikolaj. 1939/69. *Principles of phonology.* Berkeley, California: University of California Press.

Tyler, Michael D., Catherine T. Best, Alice Faber, and Andrea G.Levitt. 2014. Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica,* 71: 4-21. DOI: 10.1159/000356237

Upton, Clive, William Kretzchmar and Rafal Konopka. 2003. *Oxford dictionary of pronunciation of current English*. Oxford: Oxford University Press.

Waniek-Klimczak, Ewa. 2005. *Temporal parameters in second language speech: An applied linguistic phonetics approach*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.

Weckwerth, Jarosław. 2011. English TRAP vowel in advanced Polish learners: Variation and system typology. In *Proceedings of the 17th International Congress of Phonetic Sciences*, ed. Wai-Sum Lee and Eric Zee, 2110-2113. Hong Kong: City University of Hong Kong.

Weinreich, Uriel. 1953. *Languages in contact: Findings and problems.* The Hague: Mouton.

Wells, John C. 1962. *A study of the formants of the pure vowels of British English. MA dissertation*. University College London. http://www.phon.ucl.ac.uk/home/wells/formants/index.htm

Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.

Wells, John C. 1984. English accents in England. In *Language in the British Isles* ed. Peter Trudgill, 55-69. Cambridge: Cambridge University Press.

Werker, Janet F. 1986. The effect of multilingualism on phonetic perceptual flexibility. *Applied Psycholinguistics* 7(2): 141-155. DOI: 10.1017/S0142716400007360

Westergaard, Marit, Natalia Mitrofanova, Roksolana Mykhaylyk and Yulia Rodina. 2017. Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. *International Journal of Bilingualism* 21(6): 666-682. DOI: 10.1177/1367006916648859

Wierzchowska, Bożena. 1980. *Fonetyka i fonologia języka polskiego*. [Phonetics and fonology of the Polish language]. Wrocław: Ossolineum.

Wrembel, Magdalena. 2005. *Phonological metacompetence in the acquisition of second language phonetics*. Unpublished Ph.D. thesis. Adam Mickiewicz University, Poznań, Poland.

Wrembel, Magdalena. 2011. Metaphonetic awareness in the production of speech. In *Speaking and Instructed Foreign Language Acquisition* ed. Mirosław Pawlak and Ewa Waniek-Klimczak, 169-182. Clevedon: Multilingual Matters.

Wrembel, Magdalena. 2015. *In search of a new perspective: Cross-linguistic influence in the acquisition of third language phonology.* Poznań: Wydawnictwo Naukowe UAM.

Wrembel, Magdalena, Marta Marecka and Romana Kopečková. (under review). Speech perception in young multilinguals.

Wode, Henning. 1977. On the systematicity of L1 transfer in L2 acquisition. In *Proceedings of the Los Angeles Second Language Research Forum*, ed. Carol A. Henning, 160-169. Los Angeles, California: University of California Press.

Venkatagiri, H.S. and John M. Levis. 2007. Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness* 16(4): 263-277. DOI: 10.2167/la417.0

Zhang, Ling. 2011. Vowel length perception in Cantonese. In *Proceedings of the 17th International Congress of Phonetic Sciences*, ed. Wai-Sum Lee and Eric Zee, 2292-2295. Hong Kong: City University of Hong Kong.

Percepcja samogłosek w języku nienatywnym: oddziaływanie kategorii i cech

STRESZCZENIE

Podczas nauki drugiego języka percepcja jego dźwięków sprawia wiele trudności: oprócz rutynowego odfiltrowywania czynników środowiskowych, kategorie dźwięków pierwszego języka filtrują dźwięki drugiego języka (Trubieckoj 1939/69). Dźwięki drugiego języka, nawet jeżeli jest ich więcej albo są inne, słyszane są jako dźwięki pierwszego języka albo waga niektórych cech danego dźwięku w drugim języku jest przeceniana lub niedoceniana. Założono, że akustyczne detale danego dźwięku są jego cechami fonetycznymi, które mogą mieć znaczenie fonologiczne, jeżeli są używane do rozróżniania kategorii fonologicznych (por. Best 1995). Dotychczas uważano, że percepcja mowy w drugim języku to asymilacja tych dźwięków drugiego języka, które są podobne do dźwięków pierwszego języka lub tworzenie nowej kategorii dla dźwięków zasadniczo różniących się od kategorii języka pierwszego (Flege 1995, Best 1995, Best and Tyler 2007). Pajak i Levy (2014) sugerują, że znaczną rolę, obok kategorii, w percepcji drugiego języka odgrywają znane z pierwszego języka cechy dźwięków, których możemy użyć nawet w innych kontekstach.

Cele monografii były następujące: (1) zbadanie współdziałania kategorii i kombinacji cech w percepcji dźwięków nienatywnego języka, (2) zbadanie roli czynników wpływających na percepcję mowy w nienatywnym języku i zaproponowanie hierarchii tych cech, (3) wieloaspektowe zbadanie percepcji angielskich samogłosek przez Polaków uczących się języka angielskiego, (4) zbadanie jak związana z akwizycją nienatywnego języka reorganizacja percepcji wpływa na percepcję dźwięków języka obcego (percepcja dźwięków języka niderlandzkiego, z bardzo bogatym systemem samogłoskowym i dźwięków języka tureckiego z silnie naznaczoną kombinacją cech (samogłoska tylna niezaokrąglona), przez Polaków uczących się jednego z następujących języków: angielskiego, niderlandzkiego i francuskiego, które różnią się cechami używanymi kontrastywnie). Przyjęta hipoteza była następująca: efekty filtrowania dźwięków drugiego języka

przez kategorie pierwszego są modyfikowane przez cechy – jedna znana cecha zwiększa prawdopodobieństwo odpowiedniej percepcji dźwięku drugiego języka, a kombinacja znanych cech zwiększa prawdopodobieństwo odpowiedniej percepcji. W percepcji mowy, która ulega przekształceniom w wyniku uczenia się drugiego języka istnieje też hierarchia cech. Cechy związane z pozycją języka w jamie ustnej, na których opiera się asymilacja do kategorii pierwszego języka, zajmują w tej hierarchii najwyższe pozycje.

Percepcja samogłosek nienatywnych języków została zbadana podczas eksperymentów wykorzystujących naturalne nagrania ludzkiego głosu w językach angielskim, niderlandzkim i tureckim. Rodzaje badań obejmowały: kategorialne rozróżnianie dźwięków AXB, asymilację dźwięków nienatywnego języka do polskich kategorii samogłoskowych z pomiarem stopnia asymilacji, identyfikację dźwięków nienatywnego języka oraz porównanie podobieństwa między samogłoskami w parach. Wybór tych testów pozwolił uniknąć dylematów z wyborem zapisu sugerowanych odpowiedzi oraz porównać wyników różnych grup i wyłonić cechy, które oprócz kategorii warunkują percepcję. Badanymi byli Polacy uczący się na poziomie zaawansowanym: (1) języka angielskiego w badaniu wzdłużnym percepcji angielskich samogłosek, (2) języka angielskiego, niderlandzkiego lub francuskiego w badaniach percepcji niderlandzkich i tureckich samogłosek.

Przeanalizowano nie tylko rolę cech fonetycznych, ale też ich kombinacji, jako czynników pomocniczych, obok kategorii, w decyzjach percepcyjnych. Badania potwierdziły główną rolę kategorii, które działają na zasadzie magnesu, ale udowodniły również rolę i hierarchię pojedynczych cech w percepcji mowy w języku nienatywnym: położenie języka względem osi poziomej ma stosunkowo małe znaczenie, większy wpływ mają iloczas i zaokrąglenie ust oraz cecha mająca największy wpływ na percepcję, czyli wysokość samogłoski. W percepcji dźwięków obcego języka daje się zauważyć wpływ kombinacji cech występujących w znanych osobom badanym językach albo wpływ kombinacji cech różniących testowane dźwięki od dźwięków znanych języków. Wyniki pokazały, że percepcja nienatywnych samogłosek jest procesem bardzo złożonym, częściowo zależnym od typu zadania, z wieloma stopniowymi efektami i wynikami uzależnionymi od niewielkich akustycznych różnic pomiędzy dźwiękami.