Perceptual representations in Interlanguage Phonology: subcategorial learning in late-learners with a smaller vowel inventory

A thesis submitted to
The University of Manchester
for the degree of
Doctor of Philosophy
in the Faculty of Humanities

2018

Fernanda A. M. Barrientos Contreras

School of Arts, Languages and Cultures

CONTENTS

Final word count: 65700

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

ABX task – Discrimination task (three stimuli)
AX task – Same/Different task (two stimuli)
2IFC task – Two-interval forced choice task.

Acronyms

AOA – Age of Arrival
AOL – Age of Learning
CA – Contrastive Analysis
CE – Canadian English
CF – Canadian French
CP – Categorical Perception
CPH – Critical Period Hypothesis
GAE – General American English
GLA – Gradual Learning Algorithm
IL – Interlanguage
ISI – Interstimulus Interval
L1 – First or native language
L2 – Second language (or Foreign Language)
L2LP – Second Language Linguistic Perception model
NL – Native language
NLM – Native Language Magnet
NNS – Nonnative Speaker
NS – Native Speaker
OT – Optimality Theory
PAM – Perceptual Assimilation Model
PME – Perceptual Magnet Effect
SCA – Single-Category Assimilation
SLA – Second Language Acquisition
SLM – Speech Learning Model
SSP – Subset Principle
TCA – Two-Category Assimilation
TL – Target language
UG – Universal Grammar
UR – Underlying Representation
VOT – Voice Onset Time

ABSTRACT

In this thesis I explore the phonological nature of newly acquired perceptual representations by highly proficient late-learners of English whose L1 is Spanish, specifically in the case where two different L2 sounds are being initially mapped onto the same L1 category. I claim that these perceptual representations are not *phonemic*; rather, what these learners acquire are *phonetic* representations that can be discriminable under certain conditions in a manner similar to that of native speakers, but that are nevertheless identified as tokens of the same L1 category.

Since speech perception is a categorisation process where the acoustic input is mapped onto the *existing* phonemic categories, then late-learners will use their L1 representations when perceiving acoustic input; and will therefore have no need to create new perceptual categories. An alternative hypothesis holds that late-learners can bootstrap new perceptual categories by means of UG access, which allows them to bypass the default categorisation process and notice the difference between their L1 categories and the actual L2 input, so that new perceptual categories can be created.

This thesis focuses on the acquisition of the perceptual contrast between the open-mid back unrounded vowel /ʌ/ and the low back unrounded vowel /ɑ/, both of which are mapped onto the same L1 perceptual category /a/. Two experiments were conducted. Subjects were divided in three groups: one of native speakers (NS), a group of highly proficient nonnative speakers of English with Spanish as L1 (NNS-A), and a group of L1 Spanish speakers who were nonproficient in English (NNS-B). The experiments included identification, discrimination and rating tasks along synthesised /ɑ - ʌ/ vowel continua (7-step and 5-step).

The results showed that unlike the NS group, both groups of nonnative speakers categorised the tokens along the /ʌ - ɑ/ continuum randomly when using L2-like labels; and showing a strong preference towards /a/ when using L1-like labels. Discrimination, on the other hand, differed according to the task: discrimination of *adjacent* tokens yielded similar results across the three groups, but nonadjacent tokens showed that the NNS-A group is more sensitive than the NNS-B group and less sensitive than the NS group. Finally, prototypicality ratings showed that while NS considered the endpoints of the continuum as good exemplars of the categories /ɑ/ and /ʌ/, both groups of nonnative speakers rated all tokens along the continuum as equally good instances of /a/.

From these results I conclude that while late-learners of L2 English do not create new phonemic categories for /ɑ/ and /ʌ/, they are able to perceive a difference that nevertheless does not seem to be enough to create a category split. These findings have implications for a theory of learnability in SLA, since it suggests that late-learners have partial access to UG insofar as input alone leads to learning within the phonetic domain but not to creation of new phonemic categories.

DECLARATION

I declare that no portion of this work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT

ACKNOWLEDGEMENTS

Many life-changing episodes happened during my time as a PhD student in the UK, which undermined my mental and physical health. However, I was also lucky enough to be surrounded by a group of amazing people who were always there for me, and I would be lying if I don't say that *I got a lot of help* from them. Therefore, I'd like to thank the following people:

First of all, to my family, to whom this work is dedicated.

To my friends Laura Arman and Deepthi Gopal, who have contributed to this thesis more than they could ever imagine, mostly by providing their encouragement and company during difficult times, along with very valuable feedback on my work.

To my supervisors, especially Yuni, who was an excellent and patient guide along the entire process, and to whom I am eternally indebted. Wendell also provided important help and advice when needed.

To my peers at UoM: Eduard Artés, Carlos Gil, Glenda Lizárraga, Michaela Hejná, George Bailey, Jane Scanlon, Danielle Turton, James Murphy, Donald Morrison, James W.R. Brookes, Stephen Nichols, Pamela Morales, Sarah Mahmood, Ziyuan Pan, Ying Fan, Fang Yang, and Julia Kolkmann.

To George Walkden, who offered important career advice, moral support, and his friendship; and to Laurel MacKenzie, who was a very patient provider of stimuli for my experiments.

The Chilean community in Manchester was the main source of enthusiastic subjects for my experiments. I am also in debt with them. Special thanks to María José González and María Montt, who became my closest Chilean friends. I would also like to thank my American subjects, most of whom were contacted through Deepthi.

To my peers and professors at UdeC: Hernán Perez, Hernán León and Jaime Soto, for giving me the foundations of my work. Special thanks to my friend Mauricio Figueroa for his tech support and friendship. I also dedicate this effort to the memory of Prof. Andrés Gallardo, whose constant academic support (expressed mostly in the form of lenghty, undeserved letters of recommendation) was key in my effort to pursue a doctoral degree in Linguistics.

None of this would have been possible without the funding provided by the President's Doctoral Scholar award, for which I am very grateful. Thanks as well to the Graduate School Office at Manchester for all the help given throughout the process.

To Silke Hamann and Patrycja Strycharczuk, for their valuable feedback and comments on previous versions of this thesis.

Special thanks to Carmen Kelling and Bettina Braun at the University of Konstanz, for their constant help and efforts to make me feel at home. Many thanks as well to Alex Rehn for her company, encouragement, and friendship during the final stages of this work.

To my dear friends Lidia, Vladimir, Néstor, Constanza, Paulina, and Guido, just for being there for me when I just needed to talk.

And last but not least, thanks to Henri Kauhanen, whose help with stubborn scripts and his constant encouragement were absolutely vital in order to finish this work.

A Marianela, Iván, Consuelo, Cristóbal, Cristian,
Benita y Jorgito

CHAPTER

ONE

INTRODUCTION

Yo soy en español. En otras lenguas
me siento fatalmente como un tonto.

Teresa Eva María Rafinska
FERNANDO GONZÁLEZ-URÍZAR

The present work is an attempt to bridge the often neglected gap between the fields of Second Language Acquisition (SLA) and phonological theory, as it intends to determine the phonological nature of perceptual representations in late-learners of a second language within both an SLA and theoretical phonology framework. The task turns out to be more challenging than expected, as both areas focus on different domains of linguistic knowledge: while SLA is concerned with the interaction between two languages that have been learned sequentially (and mostly in the effects of L1 over L2), phonological theory has mainly assumed a monolingual perspective. Nevertheless, most speakers of a certain language have at least some experience in a second language no matter how little it might be, and the interaction that these two languages may have at an intra-speaker level is a fact that neither phonological theory nor SLA take into account. While SLA has often deemed the problem of acquiring a second phonological system as secondary (compared to L2 syntax and semantics), phonology assumes that the representation of sounds and the way that the grammar interacts with these representations are one and the same for all the languages learned by the speaker during their lifespan.

Such a lack of interaction between these two fields results in a theoretical gap: What do we know about the phonological knowledge acquired by speakers of an L2 that was learned at a later stage in their lives? Do these speakers acquire perceptual representations that are exactly like those of native speakers of the target language (TL), or are they somewhat different in terms of categoricity, robustness, or prototypicality?

In this work I will explore the phonological representations that late-learners of a second language create as a result of increased experience in the L2. However, we will only take into account the perceptual side of this issue, and consider the production part only to the extent that this is the perceptual input of the learner.

The specific question that this thesis explores is the process of developing two different perceptual categories for two vowels in the L2 that have been perceptually mapped onto one L1 phonemic category. The case study in question considers the way in which the /ʌ/ - /ɑ/ contrast in English is perceived by L2 speakers whose L1 is Spanish, which I claim is (a) initially mapped to the L1 category /a/; and (b) perceived as instances of this L1 category in most cases, regardless of the learner's amount of experience in the L2. Several hypotheses have been taken into account, which range from no learning to full category creation. The final aim is to move towards a conceptualization regarding perceptual category creation in L2, and how these differ from those of the L1. Specifically, this work claims that the resulting representations in cases where two vowels of the TL are mapped onto the same L1 category is not phonemic; rather, such representations are phonetic. This claim is supported by two perceptual experiments, where the performance of experienced L2 learners of English (Spanish L1) was compared against that of less experienced L2 learners as well as against that of native speakers of English.

1.1 Interlanguage phonology

The question on the specific representations acquired by L2 speakers in the phonology domain has been largely ignored, and perhaps consciously, due to the difficulty of access to these representations through experimental methods. Studies in L2 phonology are rather cautious in assigning a phonemic status to newly acquired perceptual representations, and the terminology used often refers to 'categories', without clearly stating the domain in which these categories belong. However, there is a growing body of research that uses the term 'interlanguage phonology' (e.g. Archibald, 1993; Major, 1998; Tarone, 1978), which, I sustain, is the one that best reflects the in-between nature of these categories.

What might be seen by some as a minor terminology issue is in fact a reflection of the implicit agreement that the structures of the L2 do not have the same status in the speaker's mind than those of the L1. While the L1 has a well established set of representations and structural units that have been constantly reinforced during the lifespan (and thus reaching the "final state" as posited by generativist frameworks), the L2 does not, and even more so in the case of late-learners: in fact, the lack of discussion about the final state of an L2 is rather evident (Gregg, 2001). Since an L2 is learned later in life, the L1 is the only linguistic experience that the speaker has had until the moment that L2 learning begins. If we consider a late-learner of any given L2 (that is, a person who started learning the L2 during adolescence or adulthood) then the learning process will be different due to different factors, ranging from environmental ones (such as the amount of input received and the considerably fewer opportunities to speak the L2) to more cognitive reasons, such as the use of different cognitive skills in learning. Hence, the nature of such linguistic knowledge may not be the same in many ways: processing, accuracy, speed of acquisition, and even the learning mechanisms at hand will be different.

Thus, the notion of interlanguage (IL) is crucial in L2 learning as it refers to a separate linguistic system that L2 speakers acquire during the process of learning the target language (TL) (Selinker, 1972). Therefore, an interlanguage phonology would be that of a nonnative speaker (NNS) who strives to: a) create new sound categories, both in production and perception; and (perhaps to a lesser extent) b) learn the phonological rules and constraints that regulate the output of a native speaker of the TL.

However, the concept of interlanguage should not be understood as a mere residual effect or a ‘drip’ of the L1 over the forms of the TL. Although transfer is important in the formation of the IL in all of its domains, there are other mechanisms that take place during the acquisition process, such as overgeneralisations of TL rules, or fossilization, both of which will be described in Chapter 2. But more importantly, the most powerful assumption within the IL framework is that the interaction between the L1 and the L2 can be best understood as a system in itself, with emerging patterns that can be attributable not only to transfer but also to other aspects of the learning process; and that the resultant representations have a unique configuration that does not resemble any of the structures corresponding to those found in the native language.

Despite the above, the field of L2 phonology has adopted a largely behavioral point of view, focusing mostly on NNS performance in production and perception of either a specific segment or any other suprasegmental aspect of the L2, such as stress, or tone in the case of tonal languages. Although this perspective is indeed useful as it increases awareness of what an initial state might be and which aspects of the L2 offer more difficulties to learners, there has been little interest in understanding, or even conceptualising on, the nature of these categories; or what the implications for a theory of language learning such discoveries may be.

In this regard, the main issue at stake is the lack of description for the sound representations that a nonnative speaker learns *on top of* a previously acquired system of representations, i.e. the L2 phonology over the L1 phonology. An extensive amount of research in the field of L2 phonology has proven that a default initial state is that of full transfer of the L1 inventory and phonological rules to the L2, unless the sounds of the L2 phonology are different enough to those of the L1 so that they occupy an empty space in the speaker’s phonological knowledge. But transfer is not necessarily an infelicitous situation, as redeploying elements that are similar does not constitute a problem *per se*. However, difficulties will arise for the NNS if transfer entails the loss of a given contrast in the L2 as a result of two or more elements being mapped onto only one L1 element. Nevertheless, this state tends to change over time for each case, with different levels of success as well as different implications for learnability.

This work, however, will not take into account the articulatory phenomena that occur in L2 speech production; rather, it will focus on the problem of L2 speech perception, more specifically in the case of late-learners of an L2. I will define 'late-learner' as a person who (a) has started learning a second language (L2) after having acquired a first language (L1); (b) has not learned their L2 in an immersion context (living abroad, speaking the language at home, or studied in an immersion school) before the age of 18; and (c) lived in a country where the L2 is not used on a regular basis, neither in oral nor written form.

This research will focus on two main aspects regarding the acquisition of perceptual categories in L2 when the phonemic categories in the TL are perceptually warped by the same phonemic category in the L1. The first one relates to the passage from a beginner to an advanced state of proficiency with regard to the use of their L1 phonemic vowel inventory in order to perceive L2 categories, and whether there is an increase in sensitivity in nonnative speakers across the vowel perceptual space as they gain experience in the L2. The second aspect is the nature of the perceptual categories that L2 speakers create by relating their performance in different perceptual tasks to the status of such representations in the speaker's linguistic knowledge, i.e. whether these will be categorical (or closer to the phonology domain) or more probabilistic (closer to the phonetic domain). Each of these questions has been tested with two experiments, with results that suggest that late-learners of an L2 create short-term representations of sounds that encode the phonetic differences between the L2 sounds, but are nevertheless assumed to be members of the same phonemic category.

1.2 Phonetics and phonology in perception

The field of SLA uses the term *L2 phonology* in a loose way, i.e. for acquisition of phonemic categories, phonological rules, phonetic detail, and perceptual weighting of acoustic cues, among others. Since the case study and the outlined hypotheses call for a clear distinction between phonetics and phonology, I will assume that the perceptual vowel categories that L2 learners create *should* be phonemic (as they are to native speakers of the TL) and that, as such, they pertain to the phonological domain. By *phonemic*, I will assume a mental representation with a contrastive function, which is one of the features adopted in the classic structuralist view. With regard to the perceptual side of a phonemic representation, I will consider that phonemic categories in L2 should:

(a) pass a minimal pair test;

(b) be labelled by L2 listeners in the same way (e.g. same margin of error) as native speakers would, provided that they are prototypes of the phonemic categories in question; and

(c) produce a clear perceptual boundary in discrimination tasks, which can be predicted by the phonemic boundaries as shown by a labelling task.

Any other consistent perceptual pattern will be assumed to belong to the phonetic domain. Further reasons for adapting these criteria will be identified in section 2.4.2 (Chapter 2).

However, this work will also present phenomena that belong to both domains. I will use the terms *sound system* or *phonetics/phonology* instead of *L2 phonology* to refer to the acquisition of any language-specific phenomenon at the sound level (i.e. as opposed to syntax, semantics, etc.) in the L2. These are general terms, which I use when referencing either the acquisition of small phonetic detail that can nevertheless be of phonological interest –such as aspiration in voiceless plosives– or when discussing acquisition of phonemic categories. All in all, the boundary between phonetics and phonology becomes rather blurry in SLA: what is phonetic detail in the L1 may be phonemic contrast in the L2.

1.3 The study of phonetics and phonology in SLA: scope and aims

The idea of a native sound system affecting performance in a second language is far from new. The earliest impressions about the impact of the native language onto the phonology of an L2 can be traced back to Trubetzkoy (1969), according to whom the L1 acts as a 'sieve' through which the L2 phonology is both perceived and articulated. Likewise, Polivanov (1931/1974) introduces the concept of "phonological deafness" and states that perception of nonnative sounds depends on "the complex of language habits attained by every given individual in the process of mastering his mother (native) tongue" (p.231). Such 'deafness' affects not only the perceived quality of sounds, but also the quantity. Thus, the transfer of the L1 phonological knowledge to the L2 may result in two phenomena: the presence of a foreign accent in production, and a decreased ability to perceive sounds that are not part of the native language inventory. There is in fact a vast amount of research that has studied the effects of L1 over the L2 phonological system (see Chapter 3 for a review), and also vice versa, though to a lesser extent (Antoniou, Tyler, & Best, 2012; Lord, 2008).

However, the study of phonetic and phonological effects on production and perception of an L2 is relatively new and even though it has grown considerably in the past few decades, it is still more recent than studies in other domains in the L2 such as syntax or semantics. It was not until 1995, with the development of the two most influential L2 phonology models, namely the Speech Learning Model (SLM) (Flege, 1995), and the Perceptual Assimilation Model (PAM) (Best, 1995), that the field acquired a more systematic approach and became a discipline in its own right.

Despite the increasing acceptance of these models in the field of L2 phonology, it seems that neither of them makes precise predictions regarding the final state in L2 phonological acquisition: while PAM focuses only on naïve nonnative perception, SLM is rather speculative when addressing ultimate attainment. However, a third framework, the L2 Linguistic Perception (L2LP) model (Escudero, 2005), provides a much more detailed account of L2 speech perception, according to which L2 learners create new perceptual categories via the learning of cue-weighting strategies that resemble those of the TL. This work devotes special attention to this model, since it is the most detailed and empirically-based account of ultimate attainment in L2 speech perception.

1.4 This research

This work adresses whether late-learners of an L2 with a smaller L1 vowel inventory are able to create new perceptual representations. This entails creating new categories for vowel sounds that are in contrastive relation in the L2, but are (at least initially) perceptually mapped onto just one phonemic category in the L1. More specifically, this research takes as a case study the /ɑ/ - /ʌ/ distinction in American English, where minimal pairs such as *cop* /kɑp/ and *cup* /kʌp/ are perceived by native speakers of Spanish as homophones.[1]

Furthermore, the main focus is on describing the way in which this (potentially) newly acquired phonological knowledge is encoded in the grammar. This question has triggered a few subquestions that ought to be empirically tested:

- Are the vowel categories /ɑ/ and /ʌ/ being mapped to the same L1 vowel category by native speakers of Spanish?
- Does the perceptual space in terms of sensitivity rearrange as a consequence of exposure to L2 input?
- Is one of these L2 categories perceived as being more prototypical than the other, and is that due to phonetic similarity or phonological features?

[1]The reason for choosing this variety is that the American English accent is more familiar to Latin Americans (who will be the subjects in the following experiments) than any other variety of English, mostly given the abundant input in the media. In addition, a growing presence of native speakers of English coming from the United States who work as language teachers, study Spanish, or visit for tourism, has also made direct input available for learners of English as L2 in Latin America.

- How different are the representations created by experienced L2 English learners to those of native speakers of English? Do they have a phonemic status in the L2 learner's phonological system?
- Are these representations recognised as being different regardless of the perceptual task, or are they task-sensitive?

In order to respond to these questions, two experiments were performed. While Chapters 2 and 3 are devoted to a discussion on the theoretical issues regarding speech perception in L2, Chapters 4 and 5 present and discuss the experimental findings. Finally, Chapter 6 summarises the main conclusions of this work.

1.4.1 Explaining difficulties in category creation

Chapter 2 addresses the main aspects that hinder category creation in L2. The discussion in the SLA field has been focused on three main issues: (a) the existence of a critical period, and the extent of it; (b) the effect and nature of input in acquisition; and (c) a formalisation of the critical period issue under a generativist framework, with the corresponding implications for learnability.

Secondly, difficulties in category creation in L2 have also been addressed from a psychoacoustic perspective: What are the constraints that stem from specific speech perception mechanisms? In this regard, concepts such as categorical perception (Liberman, Harris, Hoffman, & Griffith, 1957) and the Perceptual Magnet Effect (Kuhl & Iverson, 1995) have been identified as mechanisms that impede the perception of raw acoustic detail in the acoustic input, thus making category creation difficult. Section 2.4.2. will present these concepts in detail.

According to Repp (1984), "The term 'categorical' [in categorical perception] may be understood in at least three different ways, which may be called 'literal', 'phenomenal', and 'empirical"'(p. 251). Importantly, the concept of categorical perception is most often understood in the second way (i.e. as a specific phenomenon in perception of speech sounds), which "refers to the experience of discontinuity as a continuously changing series of stimuli crosses a category boundary, together with the absence of clearly perceived changes within a category" (p. 251-252). Similarly, the Perceptual Magnet Effect (PME) is a phenomenon that affects perception of vowel sounds and refers to how prototypical sounds 'warp' the perceptual space surrounding them, thus making listeners unable to discriminate between a prototype and its surrounding tokens in the acoustic space. The PME shows how listeners' discrimination is poorer when one of the stimuli in a discrimination task is a prototype (i.e. a good exemplar) of the category in question; this phenomenon has shown to be present as early in life as of 6 months old (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). Section 2.4.2.2. includes a detailed description of this phenomenon.

Even though these concepts originated in native language speech perception research, such phenomena imply that poor discrimination between sounds that the listener considers as two tokens of the *same* category would result in difficulties in creating a new perceptual category in the L2.

Finally, a strictly phonological perspective has attempted to elucidate whether perception in L2 is driven by phonetic similarity or the presence/absence of phonological features in the L1 that can be redeployed to the L2.

1.4.2 Chapter 3: L2 categories in phonology

1.4.2.1 Current models

As mentioned earlier, this work takes into account three models of L2 phonology acquisition, which are described in detail in Chapter 3: The PAM or Perceptual Assimilation Model (Best, 1995), which focuses on the perception of nonnative contrasts by naïve listeners of a non-L1 language; the SLM or Speech Learning Model (Flege, 1995), which deals with ultimate attainment in L2 and covers both perception and production; and the L2LP or L2 Linguistic Perception model (Escudero, 2005), which is focused on L2 speech perception of nonnative contrasts from a formal perspective, insofar as it offers an account of speech perception where acoustic input is mapped onto discrete categories by means of acoustic-to-phonological constraints.

Soon after the rise of the aforementioned models, an important amount of empirical research in L2 phonology has attempted to test their main claims, with the SLM as the leading model[2]. Its postulates and hypotheses were formulated as a response to the Critical Period Hypothesis (Lenneberg, 1967), according to which humans would have a maturational constraint that does not allow for language acquisition, or at least not in the way that speakers learn their native language. In this respect, the SLM takes a clear stance insofar as its main claim is that the mechanisms used in learning the L1 remain intact over the lifespan, and can be applied to L2 learning (Flege, 2003, p. 239). The other postulates derive from this main claim or complement it somehow.

PAM, on the other hand, has also inspired a research programme that looks for the bases of speech perception in nonnative speakers. This model takes a direct realist approach (Browman & Goldstein, 1989), which considers gestures (as opposed to acoustic cues) to be perceptual primitives (Best, 1995). While the direct realist approach has been criticised by researchers in the field of phonology for several reasons, such as that phonological systems show a clear tendency towards inventories with rather similar articulatory gestures (e.g. a large set of obstruents), that vowel inventories are dispersed with the goal of being easily distinguishable from

[2]Interestingly enough, the SLM makes very few predictions regarding utimate attainment. However, it has been arguably the most cited L2 phonology framework, regardless of its lack of predictive power in terms of the final state.

an acoustic perspective, and that processes of language change tend to reinforce acoustic differences (Ohala, 1996). Diehl and Kluender (1989) mention as well the fact that auditory enhancement can only make sense when taking into account the acoustic outcome of the series of articulatory gestures; this can be supported by theories such as dispersion principle in vowels, or the presence of cues that enhance each other in consonants. However, the claims made in PAM regarding perceptual outcomes in naïve speakers of a nonnative language have been extensively accepted in the field of L2 phonology.

Finally, the L2LP model has attracted increasing interest in the field of L2 spech perception, given its focus on perception (rather than production), and its exhaustive empirical approach. One of the main tenets of the model is that L2 learners need to perform two different learning tasks: a representational task (i.e. learners need to create new perceptual categories) and a perceptual task, where listeners learn a different way to deal with the acoustic signal so that it can be mapped onto the corresponding new category.

However, the link between the vast amount of knowledge regarding the structure and rules of a native language grammar and the set of representations (both in perception and production) that NNS creates is still rather weak, and none of these three models seems to address these problems satisfactorily. The SLM theorises about these representations and defines them as *phonetic categories*, thus implicitly excluding the possibility that these could be phonemic representations such as those of the native language. However, this is as far as the model goes in terms of explaining the nature of these representations, and this is where the present work attempts to make a contribution, in particular with the perceptual aspects of L2 speech. On the other hand, PAM has a strength that SLM does not have: It is able to make predictions about contrasts in L2 according to the characteristics of the L1 category inventory and how the categories of the L2 are assimilated to the former ones, thus distinguishing different types of L2-to-L1 mappings. However, this is a model that only accounts for perception of naïve speakers and does not make predictions regarding the creation of categories that take place after a certain amount of learning. Finally L2LP, while explaining in great detail the passage from initial to a final state by means of perceptual learning, it still leaves certain gaps regarding the final state, particularly in terms of whether perception in L2 can become categorical, or if it behaves in a probabilistic manner.

In sum, most of the research made in L2 phonology has considered the aforementioned models as their core and theoretical support, all of which have not directly addressed the problem of the nature of the perceptual representations that L2 speakers create, and more specifically, the way in which these representations relate to phonological theory. More specifically, this research deals with perceptual representations of segments that are contrastive in the L2, but not in the L1. This gives rise to the following research questions: are these representations of a phonemic nature, such that non-native speakers could easily recognise minimal pairs? Are these sounds just L1 categories being reused in such a way that two non-contrastive sounds in the L2 are mapped onto simply one L1 category? Would these two sounds be perceived as two different L1 categories? Would they be perceived as different when compared to each other, but keep being perceptually identified as an L1 sound? This research focuses on this theoretical gap, and aims to gather experimental evidence regarding phonological knowledge in L2 speakers.

1.4.2.2 Possible outcomes

Chapter 3 outlines four possible learning outcomes where L2 speakers deal with the perception of contrasts that are not present in the L2. The lack of such contrasts are caused by L1 sounds, which generate a warping effect for perception of L2 sounds that are acoustically similar.[3]

In Outcome 1 or **No learning**, the two L2 sounds are mapped onto the same perceptual L1 category and are perceived as the same sound. This is a complete perceptual assimilation scenario, where nothing is being created or learned. L2 speakers undergo fossilization (see definition in Chapter 3) at a very early stage and would become perceptually 'deaf' to differences in the incoming input. Hence, when minimal pairs like 'cop' and 'cup' are listened by L2 speakers who do not hold an articulatory open/mid-open height difference, nor a tense/lax contrast in their grammar (neither in /ɑ-ʌ/ nor elsewhere in their inventory), they would assume these two words to be homophones, where lexical access in these cases is mediated by lexical context. However, it is worth noting that this is only a perceptual account, as L2 speakers may have different categories in production: As pointed out by Bohn and Flege (1997), there is no correlation between these two skills, and hence the fact that certain late-learners of an L2 do produce two different target-like categories should not be taken as proof that they also perceive these categories to be different.

In Outcome 2 or **Deflection**, two L2 categories that are initially mapped onto the same L1 category start being perceived as different, resulting in one L2 category ($L2_a$) being attracted to the L1 category ($L1_a$), while the other L2 category ($L2_b$) is attracted to the following most acoustically similar L1 category ($L1_b$). With regard

[3]It is worth noting that the concept of similarity is not sufficient in order to explain the mapping phenomena in L2: As Barrios, Jiang, and Idsardi (2016) point out, the models mentioned above are relying heavily on the notion of acoustic similarity (p. 369-370). Nevertheless, this is not the only way to understand it: Other models, such as that of Brown (1998), also rely on the notion of similarity, but from a phonological point of view. In Brown's model, the notion of similarity is related to phonological features, and hence 'similar' sounds are those that share a feature. The model predicts that if the L1 grammar has the same feature contrast anywhere else (i.e. for a different pair/set of sounds), then these can be, in Archibald's (2005) words, 'redeployed' to the nonnative contrast (See Chapter 3 for a full review). I will consider the notion of 'similarity' as acoustic for the purposes of this work, as the vowel systems in Spanish and English have similar sets of features but a different set of boundaries among features (e.g. vowel openness has phonetically a three-way distinction in Spanish, versus the four-way distinction in English).

to which vowel 'stays' in the L1 category and which one 'migrates' to the neighboring L1 category, two options can be outlined: (a) one where the L2 category closest to prototypical acoustic L1 values remains, and the other is deflected to the acoustically closest L1 category; and (b) another one where warping effects created by the second L1 category is subject to language-specific cue-weighting and not necessarily overall F1/F2 values. Therefore, even an L2 sound that is not very acoustically similar can still be warped to a given L1 sound as listeners can be perceptually more sensitive to one specific perceptual cue and therefore map the L2 sound to a relatively far L1 category.

This outcome bears several implications with regard to learning new categories. First of all, it assumes that learning occurs only in the form of remapping, and that no new perceptual representations are being created: The acoustic signal of L2 vowels that were initially mapped onto the same L1 category are now perceived as different and are therefore mapped onto different L1 categories with the purpose of keeping the perceptual contrast. This would also imply that the original boundary between the L1 categories undergoes a shifting process in order to allow the input that was firstly mapped onto one of the categories to be remapped onto a different one. Hence, this would not be exactly phonological learning but mostly a movement of boundaries in the L1.

A further implication, from an experimental viewpoint, is that research on category creation in L2 that is based in recognition of minimal pairs might not address the possibility that what looks like creation of two new L2 categories is simply a mapping of these L2 categories onto existing L1 categories. This phenomenon would then call for a combined approach in experimental methods, e.g. an L2-label identification task combined with an L1-label identification task. In this regard,the first task is expected to show a consistent identification of these L2 sounds as two different categories; likewise, the second task (i.e. identification of L2 sounds as L1 categories) should yield fairly consistent categorisation results, with tokens of category $L2_a$ being mapped onto L1 category $L1_a$, and tokens of category $L2_b$ being mapped onto category $L1_b$. This second task should then test whether that it is remapping, and not category creation, that is actually taking place in the mind of the L2 speaker. For instance, if a native speaker of Spanish is asked to listen to the English vowels /i/ and /ɪ/ and label them as the vowel in either *sheep* or *ship*, then the results might look like the subject has created a new category. However, if the same speaker is asked to categorise the same stimuli as the first vowel in either *piso* ('floor') or *peso* ('weight') and the results look exactly the same as with *sheep* and *ship*, then it can be concluded that the subject has just learned to map the acoustic values for the English vowel /ɪ/ onto the Spansh category /e/. Conversely, if the outcome of these two perceptual tasks is that the subject has correctly categorised the /ɪ/ - /i/ pair with English labels but then categorises all the stimuli as instances of Spanish /i/ (i.e. instances of /ɪ/ have been categorised as Spanish /i/), then it can be safely assumed that the category has been created. Hence, it is the second task (L1-label identification) that allows for distinguishing between actual category creation from the remapping outcome.

In Outcome 3 or **Subsetting**, two subsets of a same phonemic category are created, and behave in a similar way to allophones in free variation. In a subsetting outcome the L2 learner labels the L2 sounds as members of the same L1 category, but is able to perceive the difference between them. With regard to the resulting representations, this outcome is rather similar to the concept of phonetic category as proposed by Flege (1995). However, this would not be an easy task for the L2 learner given that the phonological features involved are vowel height, and a tense/lax contrast that is not present in the NL. The L2 learner then recognises two different target mappings, but they are still within the same category; therefore, the learner would be able to perceive a difference between the two sounds, but would fail to label them as different. However, such learning challenges the findings of the Perceptual Magnet Effect (Kuhl & Iverson, 1995). According to the PME, discrimination within vowel categories is low as category prototypes have a warping effect in the perceptual space (a more detailed explanation of the concept can be found in Chapter 2, section 2.4.2.3.). *A priori*, it seems unlikely that L2 speakers will be able to retain long-term memory representations of a within-category acoustic difference.

An important implication of this outcome is that discrimination would only be available in full-L2 language mode. Language mode is a concept that refers to "the state of activation of the bilingual's languages and language processing mechanisms at a given point in time" (Grosjean, 2001, p. 3). However, such discriminatory skill need not be present in the L1; otherwise, L2 speakers would be able to perceive a difference when using the L1 as well. If subsetting takes place, then L2 speakers would have two different perceptual grammars, which would also have different degrees of sensitivity. Likewise, categories will always show an amount of overlap, as prototypical values will still be within an L1 category and will not undergo warping towards different L1 categories as in the Remapping outcome.

Finally, in Outcome 4 or **Full category split**, both acoustic mappings and perceptual representations corresponding to the copied L1 grammar are overridden by those of the L2. Thus, the L2 listener would have an L1 perceptual grammar that is used only for L1-like input, and a new perceptual grammar that is used only in the L2 and maps acoustic values onto L2-like perceptual representations. The difference between /ɑ-ʌ/ is therefore encoded by the new grammar as two different perceptual categories, each one of them with a different label that matches those of native speakers of the TL.

However, category creation by means of only input challenges two important theoretical claims. The first one is the PME, according to which the perceptual space is warped by prototypes in such a way that within-category sounds become difficult to discriminate, let alone label. A second challenge to this learning outcome is given by the Subset Principle, according to which first language acquisition takes place: When two grammars are given, the learner will choose the most restrictive one (Berwick, 1985; Dell, 1981). This, however, does not hold true for second language acquisition at later stages of development, where the effect is exactly the opposite: The L2 learner cannot see the most restrictive grammar, hence choosing the superset grammar. These two problems will be analysed in more detail in Chapter 3.

1.4.2.3 Experimental approaches

The main challenge for research on perceptual representations is the way in which the outcome of experimental approaches could be linked to a theory of representations. In this regard, several experimental methods have been proposed throughout the years, which inform on specific aspects of perception; however, these approaches are often either applicable only to the L1, or inform on *phonetic* behaviour. Chapter 3 offers a review on the most commonly used methods in L2 speech perception, and assesses their pertinence for the purposes of this work.

Therefore, and in order to tap the specific information about perceptual categories and their nature, a combined approach is needed. This work has explored a variety of experimental approaches, all of which include the use of vowel continua created by means of Klatt synthesis or source filtering using Praat (Boersma & Weenink, 2012). Procedures include discrimination between both adjacent and progressively distant stimuli in the continuum, identification of isolated sounds both with L1- and L2-like labels, and prototypicality ratings.

1.4.3 Reconfiguring the perceptual space

Experiment 1 (Chapter 4) aims to shed more light on perceptual learning in L2: How does the perceptual space change as proficiency increases? Do non-native speakers of English with a 5-vowel system in the L1 become more sensitive to the L2 vowel stimuli? This experiment aims to determine (a) whether the final state of these changes are more similar to those of the L1 grammar or the L2; (b) whether the creation of new boundaries take place, hence making room for a new perceptual category; and (c) the amount of change in boundary movement, in the case that (b) takes place.

Several 7-step continua were created: Between two adjacent English vowels (e.g. /ɑ-ʌ/), and between a Spanish and an English vowel (e.g. /o-ʌ/). Subjects were asked to perform a discrimination task between a vowel taken from the continuum and either one of the endpoints of the same continuum (e.g. between token 5 of the /ɑ-ʌ/ continuum and /ɑ/).

The part of the experiment with the /ɑ - ʌ/ continuum aims to discover whether advanced L2 speakers become able to discriminate along a continuum between vowels that are initially perceived as the same L1 segment. Results showed that more experienced L2 English learners were able to build a discrimination boundary between vowels that they initially perceived as the same, although without the same degree of accuracy as native speakers when comparing against endpoints.

The task with Spanish-English continua, on the other hand, attempts to confirm the results of Experiment 1 regarding L2-to-L1 mappings of English vowels onto Spanish vowel categories: If neither beginner nor advanced learners of English as L2 are able to discriminate between any step of the continuum versus the endpoint, then the English vowel is being perceived as a member of an L1 category, regardless of its different acoustic properties and the amount of input received. Conversely, if advanced learners of L2 English are able to consistently discriminate tokens after a certain point of the continuum against the endpoint, then perceptual learning takes place. The results showed that both groups of nonnative speakers of English do not discriminate between the vowels /ɑ/ and /ʌ/ when compared against /a/, thus confirming the initial assumption that these two English vowels are being mapped onto one category.

In sum, this first attempt in investigating changes in the perception of the /ɑ - ʌ/ contrast has led to two main conclusions. The data shows that L1 Spanish speakers of English cannot discriminate either /ɑ/ or /ʌ/ from /a/, but increased experience in English leads experienced learners to be able to discriminate between these two vowels, at least at an above chance level. This suggests that (a) There is a change from the initial to the final state; and that (b) Unlike unexperienced learners of English, experienced learners develop discrimination skills that are similar (but not completely identical) to that of native speakers of English.

1.4.4 Chapter 5: Perceiving differences vs. creating categories

One of the problems with discrimination tasks such as Experiment 1 is that they do not necessarily have to be understood as a correlate of category creation. In this regard, discrimination of within-category vowel tokens seems to be rather unlikely, particularly due to the Perceptual Magnet Effect discovered by Kuhl and Iverson (1995). Similar non-linear effects were discovered by Liberman et al. (1957), who claimed that within-category discrimination is more reduced than between-category discrimination: a phenomenon known as *categorical perception*.

Considering the above, it is now worth asking whether the behaviour in discrimination observed in Experiment 1 could be considered as evidence of category creation. Experiment 2 (Chapter 5) complements the partial findings of the previous experiment by combining three perceptual tasks: labelling, rating, and discrimination. This experiment uses synthesized vowels extracted from two 5-step continua: One between the English vowels /ɑ-ʌ/ and another between /ɛ-ʌ/. While the former is expected to trigger different behaviour in NNS than NS given, the prediction for the latter is that all subject groups should behave similarly, since both NNS and NS will perceive them as different categories. Unlike Experiment 1, this experiment tests labelling and perception over the same set of tokens and is performed on the same subjects.

Labelling is performed in both languages (i.e. Spanish and English labels), but stimuli were always tokens of English vowels. Each labelling trial is followed by a rating task, in order to establish a correlation between the place in the acoustic space occupied by the token and the categorisation. The discrimination task asks subjects to assign a 'same' or 'different' judgement to a pair of tokens that are either adjacent, or at a distance of two steps along the continuum.

This experiment sheds light over a few unsolved issues. Firstly, the experiment shows how prototypicality is related to the labelling and discrimination of adjacent tokens, and to what extent. In this respect, it was hypothesized that L1 categories act as perceptual magnets for TL sounds that are acoustically similar. This reflects on the results as a lack of discrimination between sounds that are acoustically closer to the prototype of the Spanish category. Secondly, it shows that there is no correlation between labelling and discrimination,unlike the findings of Liberman et al. (1957). And thirdly, that labelling with L2-like labels yields a random pattern in all non-native groups, whereas L1-like labels for the same stimuli trigger consistent categorisations from most of the /ɑ - ʌ/ stimuli as instances of /a/. This ultimately suggests that non-native speakers of English do not create categories for these vowel sounds, but that they are nevertheless able to perceive differences, even when the reference categories are those of the L1 (i.e. which occupy larger areas in the perceptual space).

These results suggest that the creation of *phonemic* categories in the L2 is an unlikely learning outcome. However, L2 speakers are able to create abstract, long-term memory representations that are able to encode contrasts that in the L2 are considered to be phonemic; furthermore, these learners are able to perceive phonetic detail along a rather short perceptual continuum. But more importantly, it suggests that the nature of such long-term representations is based on the perception of phonetic contrast as different but does not allow speakers to create and activate different lexical representations given a minimal pair. Furthermore, perception of such phonetic differences in the input is only available when performing specific kinds of perceptual tasks.

1.5 Notes on terminology

1.5.1 Second Language Acquisition

As Ellis (1994) points out, the term 'Second Language Acquisition' has not been exempt from controversy with regard to what SLA is and what it is not. Certain authors use the term 'second language' as opposed to a third or *n* language. I will understand 'second language' only as opposed to a 'first language': Specifically, a language that has been learned after the native one, irrespective of whether it has been learned after another additional language. This definition does not include speakers who have learned two (or more) languages simultaneously as their L1, a situation known as *simultaneous bilingualism* (De Houwer, 2009).

A second distinction has been made between a 'second language' and a 'foreign language'. Some researchers use the term 'foreign language' for cases in which the language being learned has been learned in an instructional setting, and 'second language' for languages that are spoken in the community on a regular basis, for different reasons. I will understand both terms as interchangeable, and given that the subjects of the experiments that are to be reported are from either South American countries or Spain (where English is not spoken on a regular basis), the definition will then be closer to that of 'foreign language'.

A third distinction is made by some authors (such as Krashen, 1981) between 'acquisition' and 'learning'. While the term 'acquisition' refers to an unconscious process, the concept of 'learning' entails a conscious, directed effort. I will again use both terms interchangeably. As perception and pronunciation are nowadays rarely taught in instructional settings (other than those conducing to a degree in the TL) due to the raise of communicative and task-based approaches to language learning, there is no conscious effort in acquiring the *sounds* of the L2. Hence, and as the focus is set on L2 phonetics/phonology, a narrow definition of 'learning' is not necessary.

1.5.2 L1/L2 and Native Language (NL)/Target Language (TL)

Strictly speaking, the terms L1 and L2 refer simply to the order in which two languages were acquired (Ellis, 1994, p. 12), with L1 always being the native language of a speaker, and L2 being a language that was learned later on during the life span. On the other hand, the notion of target language (TL) refers to the language that is being learnt, and is understood as opposed to the native language (NL).

SLA researchers often use L2 and TL interchangeably, although it can be argued that their focus is slightly different: while the former simply denotes the order of acquisition (that is, a language learnt *after* another), the latter may imply that the language in question is being learnt in a conscious manner. Here I use L1 and NL interchangeably, but will prefer L2 to denote 'the second language of a person', while TL will refer to 'the specific linguistic system that is being learnt' (i.e. English for the subjects of this research).

1.5.3 Notation

A distinction can be drawn between underlying representations (UR) of lexical forms and a phonemic form resulting from the application of a lexical rule. However, I use slashes for all phonemic representations in this work as this thesis does not consider cases in which derivations at an intermediate level of representation are made. Anything below the phonemic level, such as raw acoustic values, overt forms, and subphonemic categories, will appear in square brackets.

1.6 Summary

Learning the phonology of a second language (L2) is not a trivial task, neither in the case of production nor perception. Additionally, it poses different challenges for learners according to the characteristics of both the L1 and L2 phonological inventories. A late L2 learner whose L1 has a smaller phonological inventory will face the paradoxical task of learning to perceive one L1 sound as two L2 sounds. This situation leads to a crucial question for L2 phonology: Can native speakers of a language that has a smaller vowel inventory create new vowel categories as a result of a process of late second language learning? And if so, how would these categories fit into the bigger picture of phonological knowledge? Even when leaving maturational constraints aside, L1 perception plays a rather important role in perceiving the raw acoustic input in the L2. Following Trubetzkoy's sieve metaphor, in order for L2 speakers with smaller L1 inventories to achieve phonologically optimal perception in their L2, a full change of the sieve's configuration is necessary. If we consider optimal L2 speech perception as native-like perception in the TL, L2 speakers would then need to replace the relatively wide mesh in their sieve with a narrower one; a task that seems to be rather difficult if the only means through which this will be achieved is input alone. Chapter 2 will present the main constraints of category learning that L2 speakers would need to overcome.

Research on L2 phonology has developed several models of L2 speech perception, which make different predictions regarding the way in which nonnative speakers deal with categories that are not present in the L1. However, these models do not clearly address the nature of such categories; a problem that this work aims to resolve. Likewise, experimental approaches in perception also have limitations given that they can only inform on very specific aspects of perception, but not on the nature of the newly created categories. Therefore, the need for an experimental approach that takes different tasks into account is needed.

Nevertheless, several possible outcomes can be hypothesised, and it might be the case that optimal L2 speech perception does not necessarily need to equal native-like TL perception. In this respect, possibilities where L1 knowledge is optimised in order to account for gradually perceived differences in the L2 are also part of the spectrum of possible outcomes in perceptual learning.

Two experiments will be presented in order to understand the way in which L2 speakers use their native phonological knowledge, which are reported in Chapters 4 and 5. The results lead to the conclusion that the creation of phonemic perceptual categories by late-learners of L2 English with a smaller native vowel inventory is unlikely, but that further experience in the L2 leads to phonetic learning that allows advanced learners to perceive differences in discrimination tasks. This outcome has implications for a theory of L2 learnability, as it confirms previous theoretical formulations regarding the unlikeliness of native-like performance in late-learners.

CHAPTER

TWO

CONSTRAINTS ON CATEGORY ACQUISITION IN INTERLANGUAGE PHONOLOGY

Tengo una repulsión innata por todo lo gelatinoso, sobre todo, si es del color llamado vulgarmente concho de vino. De más creo agregar, que si ello es de tal color y tiene patas, la repulsión me es completa y, por lo tanto, intolerable. Bien podría ser que tras el sofá no hubiese tal cosa, pero bien podría ser también que la hubiese.

Ayer
Juan Emar

2.1 Introduction

Why is the perception of phonemic contrasts in the L2 difficult for late-learners? In this regard, it is important to first draw a distinction between learning structures such as syntax or morphology, and learning the sound system of a given language. While SLA theories largely agree that all types of linguistic structures in the L1 will somehow transfer to the L2, it seems that L2 learners rarely achieve native-like performance in the realm of phonetics and phonology (Ellis, 1994, p. 491–493).

This distinction entails that along with a baseline level of difficulty that is transversal to all linguistic domains, a second layer of difficulty should be taken into account, which is given by the nature of the mechanisms involved in speech perception and the way in which the mind deals with acoustic information when processing language.

This work illustrates the extent of these constraints in category acquisition by taking into account a specific case study: Perception of the /ɑ/ - /ʌ/ contrast by L1 Spanish speakers of English as L2. This chapter will first describe the difficulties in acquiring this contrast from a phonological perspective, and will then proceed to list and summarise the main reasons why the acquisition of non-native contrasts in L2 phonology presents difficulties for the L2 learner from three different perspectives: linguistic, maturational, and psychoacoustic. These three streams of discussion are fundamental in order to understand the factors that hinder the acquisition of perceptual contrasts.

2.2 Linguistic constraints

2.2.1 Transfer

One of the most important factors that should be taken into account in the acquisition of L2 categories is the relationship between the L1 and the L2 in terms of the differences in the inventory: SLA theory and a vast amount of research in the field of interlanguge phonology has shown the proponderance for transfer processes in the rise of IL structures.

Transfer has been defined as "the incorporation of features of the L1 into the knowledge systems of the L2 which the learner is trying to build" (Ellis, 1994, p. 28) and has its origin in behaviourism, according to which language learning was no different than any other type of learning: Both L1 and L2 learning were mostly seen as a matter of habit formation through repetitive stimuli and the corresponding reward when a target-like response was given, or correction after non-target responses (Ellis, 1994, p. 299).

The notion of transfer has a largely behavioural basis and was extensively studied by the founder of Contrastive Analysis (CA) Robert Lado (1957), according to whom the only way to understand L2 errors is to compare the features of the L2 to those of the L1. A strong version of CA claims that *every* error made in the L2 is caused by interference from the L1. However, a weaker version of CA was proposed later, where the L1 is seen as a useful tool to understand *some* IL errors but is by no means the only source of non target-like linguistic forms. In accordance with this view, CA mostly presented SLA as a matter of transfer from the L1, where a comparison between the NL and the TL could become a predictor of the most frequent errors that language students would make and which aspects of a specific L2 would be more difficult to grasp. Transfer could then take two different forms: negative transfer (L1 $\neq$ L2, which results in *errors*) or positive transfer (L1 $=$ L2, or *facilitation*).

As behaviourism became less popular in language studies and other approaches such as generativism gained acceptance in linguistics, this L1-determined understanding of errors in SLA lost its explanatory power. More importantly, as studies in L2 errors increased, cases where L1 interference was not the cause of L2 errors evidenced that a new theory was needed in order to account for such phenomena.

It was in light of these facts that Selinker (1972) coined the term *interlanguage*: A linguistic system in its own right, that differed from both the L1 and the L2 and showed linguistic forms that could not be accounted for by just L1 transfer and target-like L2 forms. Selinker considers five psycholinguistic processes as possible sources of IL forms, where transfer is just one of them: overgeneralization of TL rules, transfer of training, strategies of communication and strategies of learning.

Nowadays, CA has been largely disproved due to the increasing evidence which shows how IL forms can present errors that do not originate from the NL, and that are better explained by the processes described by Selinker. More current approaches in SLA understand transfer as a cognitive strategy instead of a habit, and can be described as "the influence resulting from the similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired" (Odlin, 1987, p.27, as cited in Ellis, 1994, p. 301).

2.2.2 Interlanguage at the phonetics/phonology level

While the list of processes identified by Selinker are comprehensive descriptors of the second language learning process, not all of them apply to the phonetics/phonology domain. Furthermore, transfer is still the most evident mechanism that causes non-target-like structures in the learner's IL at the sound level. Such structures may include transfer of the elements of the phonological inventory and the articulatory and acoustic specifications for a specific sound; note for instance, the differences in production of /æ/ and /ɛ/ in Canadian English and Canadian French (Escudero & Polka, 2003); as well as the transfer of phonological rules, such as /e/ epenthesis before English /sC/ clusters by L1 Spanish learners (Carlisle, 1988). However, most of the interlanguage research comes from the syntax and morphology domains; very few specific claims have been made from phonology (see Archibald, 2005, and Broselow & Finer, 1991, for two exceptions related to suprasegmental phenomena).

One of the main assumptions of this work is that the L2 learner's initial state is characterised by a full transfer of the L1 inventory, which in the case of L1 Spanish speakers of English would entail the lack of contrast between /ʌ/ and /ɑ/, given the smaller number of elements in the Spanish vowel inventory. In our case study, one of the elements of the L1 inventory, /a/, becomes the target of a perceptual mapping from two vowels in the TL. This process of double-mapping seems to pose a major problem for the learner, since moving away from these transfered representations is constrained by mechanisms that are specific to speech perception (see section 2.4).

In this respect (and probably only in the case of acquisition of L2 sound systems), the CAH can be regarded as a valid theory with predictive power, although it does not necessarily account for cases where a certain L2 phonemic category does not form part of the L1 and complies with two conditions pointed out by Flege (2003) and

Best (1995): That the category in the TL is sufficiently dissimilar and salient enough when compared against the categories in the NL; and that the perceptual cues are salient so as to make them easily perceivable. Again, phonemic representations pose a new challenge for SLA theory: one size does not fit all.

However, assuming a full transfer hypothesis for the purposes of this work does not rule out the fact that the resulting representations would be better described as interlanguage representations, rather than mere replications of L1 representations. As mentioned in Chapter 1, this work uses the term *Interlanguage Phonology* in an attempt to visibilise the idea that the nature and quality of the linguistic knowledge obtained during learning a second language phonology may be neither that of the native language nor the target language. In fact, the experiments presented in Chapters 4 and 5 will show that the resulting representations, although originally formed via L1 transfer, evolve into a specific type of representation that does not bear resemblance neither with the L1 structures nor with those of the L2. In this regard, the concept is particularly descriptive of the *type of cognitive representation* to be described: The perceptual representations acquired through a learning process are characterised by an initial state consisting of a full linguistic system, and an undetermined final state that aims to resemble the target language, but does not necessarily become a calque of the TL grammar.

More importantly, IL as a concept makes a claim regarding the 'final' state of language learning, insofar as it emphasises the fact that L2 learners develop a systems that is not the TL itself, regardless of their stage. It is also worth noting that the concept refers to a notion of competence, and not performance. While an L2 speaker might develop target-like fluency, it does not follow that the language processing strategies or the system of representations of the L2 learner are the same as those of native speakers. In a simple analogy, one could think of a mathematical operation in a second grade classroom, such as 5 x 3. While some students found the answer by simply having memorised their times table, others could have slowly worked it out by calculating 3 + 3 + 3 + 3 + 3; others perhaps tried 5 + 5 + 5. In the end, all students were able to find the right answer, but their cognitive strategies and representations differed. Furthermore, this analogy could extend to the way in which IL errors can hint at the strategy used: If a student writes down '12' as the result, it is very likely that the student used the 3 + 3 + 3 + 3 + 3 strategy and forgot to add the last 3 to the final sum. Likewise, it could also be the case that the student gave the right answer by calculating 3 + 3 + 3 + 3 + 3, and the teacher will never know that the student has not memorised her times table.

2.2.3 The nature of the evidence: input in perception

In addition to the constraints given by transfer, speech perception in L2 sets very specific difficulties for L2 learning. In this regard, there is little discussion on the nature of input in L2 speech perception: What could exactly count as negative evidence in perception? Since perception is a 'hidden' process, then how can L2 speakers receive feedback? Production errors, on the other hand, are more likely to be corrected, but correcting perceptual mergers is much less straightforward. This section refers to the role of input in perception.

2.2.3.1 Input and intake

In SLA, input is usually understood as anything that the learner reads or listens to in the TL. Smith (1993) defines input in a broad sense as "language data that the learner is exposed to, that is, the learner's experience of the target language in all its various manifestations" (p. 166). However, Corder (1967) acknowledges the fact that not everything that the learner is exposed to may result in actual learning: "The simple fact of presenting a certain linguistic form to a learner [...] does not necessarily qualify it for the status of input, for the reason that input is 'what goes in' not what is *available* for going in, and we may reasonably suppose that it is the learner who controls this input, or more properly his intake" (p. 165). This view of input as a concept that does not necessarily refer to acquisition of the forms that are present in it gave birth to the concept of *intake*.

> Gass and Selinker (1994) view intake as "the process of assimilating linguistic material" (p. 302). It is "the mental activity that mediates between input and grammar and is different from apperception or comprehension as the latter two do not necessarily lead to grammar formation" (p. 302). This, as they say, suggests that 'intake' is not merely a subset of input. Rather, 'intake' and 'input' refer to "two different phenomena" (p. 302) (Ying, 1995, p. 182).

How does this relate to the acquisition of a sound system in the L2? The input-intake distinction is still valid in this realm insofar as acquiring a new sound inventory also involves an apperceived input (i.e. the sounds of the L2) that needs to be internalised in order to become part of the L2 learner's linguistic knowledge. However, research has largely proven that even speakers who received large amounts of input (i.e. lived in a TL speaking country for many years, are fluent in the TL, etc.) are still much less accurate in perception than production (e.g. Bohn & Flege, 1997) There seems to be a missing bit in the equation: How do we move from just listening to the input to the actual intake of these perceived sounds?

One possible answer for this question relies on the concept of *noticing*, which is defined as a process that involves both awareness and consciousness, thus giving the concept of attention an important role in learning (Gass, 1997). However, something needs to happen in order for this noticing process to take place. Some SLA research has suggested that the key to achieve noticing and therefore intake is *enhanced* *input*, or "input that can be enhanced by an external source (e.g., a teacher) or an internal source (learners relying on their own resources)". (Gass & Selinker, 2001, p. 320). More particularly, it refers to form-focused instruction, which is a "meaning-focused activity into which an attention to form is embedded" (Gass & Selinker, 2001, p. 320). In this regard, enhanced input is usually targeted to the acquisition of syntactic forms, which are usually overlooked by the L2 learner as meaning takes precedence over form in L2 language processing. If that is the case even for syntactic forms, it is even less likely that the learner would devote some of their processing resources to the perception of detail in speech sounds:

> I take as a point of departure the following claims: that during interaction in the L2 (1) learners are focused primarily on the extraction of meaning from the input [...], (2) that learners must somehow "notice" things in the input for acquisition to happen [...], and that (3) noticing is constrained by working memory limitations regarding the amount of information they can hold and process during on line (or real time) computation of sentences during comprehension (VanPatten, 2004, p. 7)

VanPatten's quote illustrates the way attention is prioritised in a multi-level phenomenon such as language. Since meaning is the ultimate goal of linguistic interaction, the attentional resources focus on that aspect, rather than on the actual form or substance of the utterances. If such is the case, then phonetic/phonological processing is the least important in the list of priorities of the L2 learner in terms of attention; hence, as long as something is perceived and it works as a vehicle to the final meaning, then the acquisition of a target-like sound system stops, thus reaching a state of fossilization.

At this point a question arises: What stops late-learners of an L2 from *noticing* the input/output mismatch so that intake of new L2 sounds takes place? Here two possible options should be considered. The first one relates to evidence: As soon as an L2 speaker acquires a certain level of competence so that all the communicative exchanges happen successfully on a regular basis, and no corrections or misunderstandings take place, then the learner assumes that no further action needs to be taken; no negative evidence leads to a stabilisation of the system. The second option is that the learner somehow notices that *something is wrong.* This may lead at the same time to two different cases: (a) that the learner notices what specific aspects of her output do not match the input and therefore, in a need to correct her pronunciation, she starts paying further attention to the acoustic aspects in the input that differ from what she produces. This would be an optimal situation in which the learner enters phonetic perception mode, as defined by the ASL model (Strange, 2011), this model will be further explained in Section 2.4.2.1. On the other hand, option (b) is much less optimal: While the learner has already noticed that her accent is nonnative, she simply *cannot* perceive the specific acoustic detail that makes her sound nonnative; that is, the learner has noticed a certain mismatch, but fails to notice the exact aspects in her pronunciation that create this situation. Yet, while these scenarios are possible for production, they do not have any obligatory implications for perception.

Additionally, research in L2 speech perception and loanword phonology has shown the effect of orthographic input (e.g. Escudero & Wanrooij, 2010; Bassetti, 2008; and for a formal model, Hamann & Colombo, 2017), where the grapheme-to-phoneme mappings of the L1 are transfered to L2 written input. In this regard, it is also worth considering the possibility that since English and Spanish use the same Latin script despite the wide difference in the phonemic inventory, L2 speakers simply assume that the L2 has the same sound inventory as the L1; hence, L2 learners who receive more written than aural input have no reason to suspect that the English vowel inventory is at least twice as large as the Spanish one.

2.2.3.2 The nature of perceptual input

While the importance of input has been stated by every theory of language learning, and the field of Applied Linguistics has carried out a great deal of research in an attempt to find the 'holy grail' of input, (that is, the type of input that would lead to intake of grammatical forms in the TL and avoid fossilization of non target-like IL forms), there is a conceptual gap regarding the nature of input in perception. So far, we have been assuming that the only input in perception is normal speech production by a native speaker of the TL, which is the expected type of input in a more naturalistic setting. However, it seems that this is not sufficient in order to generate target-like perception in L2 speakers.

Importantly, in a competence-based view of proficiency this could go largely unnoticed. One could think of an L2 speaker of English whose L1 is Spanish in the context of a party. This L2 speaker could listen to an utterance such as "the cups are in the cupboard", and diligently reach for the cups and bring them to the table. What remains unnoticed is that the word *cups* triggered a non-targetlike underlying representation: /kaps/, which is the same one triggered hours later by a different utterance: "The *cops* are outside and are asking us to turn the music off". However, our L2 speaker is certain that no set of cups has become sentient, and that there is a couple of policemen who were sent by an annoyed neighbour. This takes us back to the second grade classroom and the student who calculated 3 + 3 + 3 + 3 + 3 = 15.

From this point on, there is little room to hope for this L2 speaker to create a more native-like type of underlying representations for the words *cop* and *cup*, unless an ambiguous semantic context is presented such as one where the likelihood of refering to a vessel that contains drinks is the same as that refering to a policeman. And even if this were the case and confusion and ensuing clarification takes place, there is still the case that this L2 speaker may not be able to reset her perceptual representations; in fact, she might even remain *unaware* that there is a difference.

While SLA does not reach an agreement on what exactly an 'end state' is, *fossilization* is one of the possible answers. Fossilization is defined as "linguistic items, rules and subsystems which speakers of a particular NL will tend to keep in their IL relative to a particular TL" (Selinker 1972, as cited in Ellis, 1994, p. 353). Although these forms may disappear, they tend to reappear in casual speech and are assumed to be caused by a combination of internal and external factors such as age, lack of desire to acculturate, communicative pressure, lack of learning opportunity, and nature of the feedback on the learner's use of the L2 (Ellis, 1994, p. 353-354). In the case of acquisition of new sounds in the L2, fossilization can be observed in the form of a permanent L1 category transfer, which can be assumed to become worse given that the mechanisms involved in speech perception are specifically designed to deal with a wide range of variation.

Thus, if the perceptual input given is only native-like production in everyday communication, then it is unlikely that changes in the perception of L2 phonemic categories that were mapped to the same L1 category during the initial state will ever take place. Since there is no urgent communicational need to change perception (or at least, so long as this is what the L2 learner assumes), the most likely scenario under these conditions is that of fossilization, where both TL underlying representations /kʌp/ and /kɑp/ remain as /kap/ in the L2 speaker's IL. If this is the case, then one should think of a different kind of input and evidence, which would then have to be a very specific type of corrective feedback that does not occur in a naturalistic setting and take the form of minimal pairs and/or isolated vowel drills. However, such specific input and corrective feedback would lead to a type of learning where the cognitive mechanisms in action are not the same as those used to learn the L1, which takes us back to the question of whether age imposes a constraint on the type of available learning mechanisms; such constraints will be mentioned in the following section.

2.3 Maturational constraints

The term 'Maturational constraints' refers to the effect of age in learning an L2. It has been largely assumed that 'the earlier the better', a claim which has been difficult to prove mainly due to (a) the myriad of factors affecting the learning process; and (b) the difference in methods that have been used by SLA scholars in their studies. Moreover, studying a complex cognitive process such as learning a second language through an experimental approach is rather difficult as some variables are impossible to isolate or quantify.

The field of SLA has considered different approaches to the age question. On the one hand, we have a merely descriptive approach that stems from biology, which is better known as the Critical Period Hypothesis (Lenneberg, 1967). On the other hand, more formal approaches have also been considered, which stem mostly from a generative framework. This section briefly summarises both approaches.

2.3.1 The Critical Period Hypothesis

The Critical Period Hypothesis (CPH) has been one of the most controversial claims in SLA, as the evidence obtained is inconclusive. The CPH states that success in learning a second language depends on whether the process takes place within a certain year span, which has usually been identified as childhood and the onset of puberty (Ellis, 1994, p. 484). According to Singleton (2005), the CPH started as a general theory of learning and attempted to account for certain biological processes in different species, which occur within a time frame that is considered as critical and after which learning will not occur. Not too long after, the theory was applied to language learning by Lenneberg (1967), who considered age two and puberty as the onset and offset of language learning, respectively. While the hypothesis was originally stated for L1 learning, it became applicable to SLA as well.

However, there is a rather large amount of variation in what SLA scholars have understood as CPH. In this regard, Birdsong and Molis (2001) point out that regardless of these differences, all formulations share a basic agreement on two main points, namely, that "learning during a critical period is assured, similar across individuals, normatively described, and probably governed primarily by endogenous factors" and that "learning outside the critical period is different in both form and success, especially in that it would be less certain and more erratic in its outcomes" (Bialystok & Hakuta, 1999, p. 164, as cited in Birdsong & Molis, 2001, p. 236).

More importantly, the evidence regarding the existence of a critical period in SLA is not just inconclusive, but it also restricts mostly to the morphology and syntax domains. In their often-cited study, Johnson and Newport (1989) test the ability to make grammaticality judgements in non-native speakers of English whose L1 was either Chinese or Korean, and find that proficiency correlates negatively with the age of arrival (AOA), hence supporting the hypothesis of a critical period affecting linguistic competence in the L2. However, Bialystok and Hakuta (1999) while obtaining similar results with data extracted from a census carried out in New York City, do not rule out the possibility of confounds arising from this type of pattern (the language-specificity of the experimental tasks, the L1/L2 linguistic difference, and the educational level of the subjects were only a few of the concerns). Moreover, they point out that data showing an clearly inverse correlation between AOA and proficiency does not necessarily count as evidence of a critical period: If this were the case, then one should expect a decline of the regression line only *after* a certain age, and not one starting from zero as in the data analyzed. In his review, Ellis (1994) considers several studies and even further replications, but the evidence is again controversial, at least with regard to grammaticality judgements (p. 484-489).

How does the CPH apply to acquisition of L2 sounds? It seems that evidence regarding production and perception is less divergent. In this regard, Young-Scholten (1994) points out that the CPH has been somehow sustained by the fact that few adults acquire native-like L2 sounds either in production and perception, although this does not rule out the possibility of native-like performance; rather, the reason seems to be the type of input in L2, which needs to include negative evidence as well.

From a production perspective, Bongaerts, van Summeren, Planken, and Schils (1997) suggest,on the basis of the reports of a study on Dutch learners of English, that native-like pronunciation at a later age is not *impossible*, but neither is it *frequent*. The reason for the difference in levels of attainment seems to be a combination of factors that make native-like pronunciation more (or less) likely to occur, namely: learner, learning context, and language variables.

On a similar note, Piske, MacKay, and Flege (2001) offer a detailed account of foreign accent and ultimate attainment in L2, concluding that the research carried out to date "does not disprove the existence of a critical or a sensitive period for L2 speech learning. They rather support the finding of previous research that AOL is the single most important predictor of degree of L2 foreign accent" (p. 212). However, this does not necessarily entail that there is a critical period in the sense of an open time window that suddenly closes; rather, the effect is linear, which also opens the possibility for more factors affecting foreign accent.

In sum, there is no consensus on the extent of the validity of the CPH, given that empirical research has proven it to be a rather multidimensional phenomenon. In this respect, Singleton (2005) points out that so far, the CPH "as it stands it is like the mythical hydra, whose multiplicity of heads and capacity to produce new heads rendered it impossible to deal with" (p. 280). In fact, the most cited evidence supporting the CPH (namely, that late bilinguals do not seem to master the L2 in the same way as their L1) obscures a number of facts that current research has not been able to account for. AOL has been one of the most used predictors for L2 mastery, but other variables such as amount and quality of input received, along with other individual differences such as motivation, language aptitude, etc. cannot be always objectively quantified and might be playing an important role in the degree of ultimate attainment. However, it is worth noting that an important distinction must be made between understanding CPH in absolute terms (i.e. whether late-learners are completely unable to learn certain structures in the L2) or if its extent is only limited to native-like attainment; distinction that is, as will be demonstrated in the following chapters, crucial for a description of L2 representations in phonology.

2.3.2 Access to Universal Grammar (UG)

The discussion regarding the role of age in L2 learning has also drawn the attention of the generativist framework. In this respect, generativist approaches to SLA have operationalised this apparent lack of ability to learn an L2 in an L1-like manner (both in terms of success and the learning processes involved) during adulthood through the problem of access to Universal Grammar (UG).

In a generativist approach to SLA, L2 acquisition by late-learners is explained by the degree of access to UG after the L1 has already been acquired, and if such a degree of access is exactly the same as when L1 acquisition takes place.

There are diverging views within the generativist framework about the type of access that L2 learners may have to UG. On the one hand, most SLA researchers agree that late-learners of L2 rarely show native-like proficiency; such a fact is usually considered as evidence for either a partial-access or no-access hypotheses. According to this viewpoint, learning in the L1 is possible by means of mechanisms that cannot be replicated by the L2 learner (White, 2003, p. 15-16). This standpoint is better known as the Fundamental Difference Hypothesis (Bley-Vroman, 1989), according to which the difference between L1 and L2 learning is deemed *internal* (caused by a different cognitive state of the adult vs. that of the child), *linguistic* (caused by changes in the language faculty, not learning in general), and *qualitative* (the domain-specific acquisition system is not attenuated, but entirely unavailable) (p. 50).

On the other hand, native-like competence, although rarely attested (particularly in late-learners), is possible, and has been seen as evidence for the full-access hypothesis. The Parameter-setting Model (Flynn, 1987) assumes that linguistic parameters (such as Pro-Drop or Null Subject in syntax) have a certain setting ('+' or '-') and that the corresponding value reflects a property of the language. Hence, a language with a '+' Pro-Drop parameter setting will allow sentences without an explicit subject: For instance, in a Pro-Drop language such as Spanish, a sentence like *Caminamos* (walked-1.PL, 'We walked') is grammatical, while **Walked* in English (which has a '-' value for Pro-Drop) is not. The L2 learner then acquires the TL grammar by either copying the parameter values in the L1, or changing them as they discover from the input that the TL has a different configuration. In addition, the Full Transfer/Full Access Hypothesis (Schwartz & Sprouse, 1996), claims that the L1 serves as a *bridge* between the initial state and UG, as incoming L2 input cannot be processed by the L1 grammar which therefore forces its restructuring (p. 41).

Access to UG in SLA has also been addressed in regard to the Subset Principle (SSP) (Berwick, 1985; Dell, 1981). Originally conceived as a first language acquisition theory, it promptly made its appearance in SLA under the assumption that the principle would not apply to L2 learning. The Subset Principle states that when first language acquisition takes place, the learner will choose the most restrictive of two possible grammars, even though the type of evidence given is mostly positive. This, however, does not hold true for second language acquisition at later stages of development, where the effect is exactly the opposite: L2 speakers choose the least restrictive grammar. The SSP explains why a child's utterances while learning the L1 are ungrammatical only to a small extent, and why this is not the case for late-learners of an L2. In contrast, if an L2 speaker is learning a TL whose grammar is more restrictive than that of the L1, then the positive evidence received will not be enough for them to correct any ungrammatical output in the L2 that derives from the transfer of the superset parameters. Hence, it has been stated that adults need nore than just positive evidence in order to acquire an L2: Negative evidence in the manner of corrective feedback seems to be necessary in order to access UG (Young-Scholten, 1994; Carroll, 2001).

Acquiring a phonemic inventory in a given L2 with more elements than those found in the L1 can be seen under the perspective of the SSP. If we assume a model of vowel representations where the F1/F2 acoustic space is divided by certain acoustic parameters, then we can assume that the area delimited by the points[1] (F1=300, F2=2500), (F1=600, F2=2400), (F1=600, F2=1800), and (F1=300, F2=1800) has three categories for L1 English speakers (/i/, /e/, and /ɪ/), whereas that same area has two categories for L1 Spanish speakers (/i/ and /e/). These categories are delimited by boundaries, which can be understood as acoustic-to-phonological constraints (i.e. cue constraints) which are imposed by the grammar of each language, as proposed by Boersma (2009) and Escudero and Boersma (2004). Different types of cue constraints have been proposed: while the first formulations were of the type *Warp (F1: [440], /300/): "do not initially classify an acoustic input of 440 Hz as a high vowel" (Boersma, 1998, p.147), subsequent versions map the acoustic input directly to the segment, such as in "F1: [300] not /i/: an F1 of 300 Hz is not /i/" (Escudero, 2005, p. 48).

Now, it could be assumed that more boundaries in the acoustic space imply a more restrictive grammar that maps acoustic values onto many different phonemic categories. Likewise, a language with fewer elements is governed by fewer rules, since the space needs fewer boundaries. Hence, Spanish can be understood as having a less restrictive perceptual grammar, thus being the superset in the English/Spanish pair[2].

With this model in mind, it can be easily seen how L2 speakers could fail to process the acoustic input, since it only comes in the form of positive evidence. As Young-Scholten (1994) points out, "only negative evidence would be able to inform

[1]Values in Hertz, after the study by Bradlow (1995).

[2]Likewise, one could also think of a feature-based model with similar characteristics, where rules derive vowel features (i.e. "the space within these points should be parsed as [+high, -low, -tense]". Here, the subset grammar becomes immediately more restrictive, as it includes an extra feature [±tense].

the learner that L1-based assumptions will result in ungrammaticality in the L2" (p. 200). This assumption is crucial with regard to speech perception, mainly due to the type of input received: What can be considered as perceptual input? And perhaps more importantly: How does the learner deal with the input? Does input become intake? What are the limitations on evidence processing? Section 2.4 will refer to this issue.

2.4 Psychoacoustic constraints: Perceptual mechanisms

While the previous section has outlined constraints that could hinder the perception of non-native structures in the language in general, one could still wonder whether the acquisition of phonemic contrasts in the L2 is actually different than acquisition of structures from other domains such as syntax or morphology. In fact, and as it was mentioned above, maturational and input-related constraints could affect any aspect of L2 learning below the level of meaning, especially considering that attentional processes are focused on the latter.

However, and as mentioned earlier, acquiring phonetic/phonological knowledge in an L2 at a later stage in life is constrained by the specific mechanisms of speech perception, which make categorisation and normalisation of the speech signal possible.

As for the scope of this work, I have already stated that late-learners of L2 English with a smaller vowel inventory could fail to perceive the contrast between vowels that occupy the acoustic space of one single vowel in the L1. This is due to the fact that perception across the acoustic space is non-linear: Within-category tokens can be perceived as being the same regardless of their acoustic distance, and noticing differences between tokens of the same phonemic category is possible only under very specific conditions. On the other hand, perception of speech sounds is strongly biased by the L1, thus making it difficult for L2 speakers to a) notice the difference, and b) create new categories for L2 sounds that could perceived as instances of the same L1 category. Most L2 speech perception models take this phenomenon into account, and a great deal of empirical research has shown how these two aspects of speech perception have hindered the acquisition of non-native contrasts in L2 speakers. Empirical approaches to speech perception in NL can be traced back to the late 50's with the Haskins Laboratories research programme, which focused on the perception of speech sounds using acoustic continua. Later research in speech perception such as the NLM (Kuhl & Iverson, 1995) seems to be crucial in this matter, as it concludes that perception is somehow biased by linguistic experience as early on as of 6 months old, although this is by no means proof that humans are unable to perceive the raw acoustic input after this period; rather, it seems that language-specific perceptual mechanisms choose not to perceive some of the acoustic details in order to make language processing more efficient.

2.4.1 Describing L2 speech perception

In order to move towards a description of L2 speech perception, we will first need to make a few assumptions regarding speech perception as a process, and the aspects of it that are of interest. First of all, it is important to understand perception as a **categorisation** process, where (a) raw acoustic input is received; and (b) this input is mapped onto perceptual categories (Holt & Lotto, 2010; Boersma, 2009). In this regard, the task that best reflects this process is an identification task (Boersma, 2009, p. 65). We will then advocate for a speech perception model that is language-specific, insofar as it takes the acoustic input and maps it onto language-specific phonological representations, like those proposed by Boersma (1998, 2009) and Apoussidou (2007). In this approach to speech perception, the listener takes the acoustic input and maps it onto perceptual representations. These perceptual representations are created by the listener during the L1 acquisition process as she receives the input. In this case, the input is tokens of speech sounds with acoustic values that obey to a normal distribution along a certain acoustic dimension. Such a distribution determines the mean acoustic values of a category and the dispersion of the range of values, which in turn will become a more abstract category. Once these categories are created, listeners categorise the incoming input following a probabilistic criterion. Thus, perceptual categories and their boundaries are formed on the basis of linguistic experience.

However, L2 speech perception is a different process given that: (a) the entire categorisation process presents additional noise given by a different system of representations; (b) the input does not always match the prototypical perceptual targets of the L1; and (c) the phonemic representations intended by the source are not necessarily present in the listener's inventory. Thus, an optimal L2 speech perception process entails *creating* new representations when the L1 does not provide them (Escudero, 2005). In this regard, a complete description of L2 speech perception also needs to focus on the acquisition of these new perceptual categories, which might not take the final discrete form that L1 categories have.

2.4.2 What makes L2 speech perception so different?

L2 speech perception is different from native speech perception in many respects. First of all, the L2 is acquired (especially in the case of late-learners) after the L1 has reached its final state. This situation not only leads to a difference in the circumstances of the learning process, but also to a qualitative difference in terms of the learning mechanisms involved: According to the Fundamental Difference Hypothesis, while the L1 is learnt via UG, the L2 is not (Bley-Vroman, 1989). As discussed in section 2.3.2, this is just one of the possible ways to describe the L2 learning process in terms of access to UG, with Full Access (Schwartz & Sprouse, 1996) and Partial Access (Schachter, 1988) as other possible options.

Additionally, L2 learners faces a different set of challenges regarding the acquisition of a new sound inventory. Perception in particular is constrained by specific mechanisms that are considered innate to humans and from which listeners are unable to opt out. In this regard, the research programme carried out by the Haskins Laboratories (Liberman et al., 1957 and subsequent) on the perception of phonemic categories have shown that within-category discrimination between tokens at an acoustic distance X is more difficult than across-category discrimination between tokens at that same distance X. This phenomenon has been known as *Categorical Perception*, although the term has been used in a rather loose way, ranging from the literal meaning to specific empirical meanings (Repp, 1984). Furthermore, the PME (Kuhl & Iverson, 1995) shows how the L1 phonemic categories create a warping effect in the perceptual space, thus making listeners unable to perceive within-category differences when the tokens of a category are close to the prototype.

These phenomena lead to an important difference between the acquisition of categories in the phonetics and phonology levels versus acquisition in other domains of language. Perception of phonemic categories is mediated by psychoacoustic processes that optimise the perception of sounds as members of L1 categories, thus hindering acquisition of new L2 categories. Since the perception process is strongly biased towards perceiving *something*, rather than not perceiving anything at all or creating new categories for every sound that seems slightly deviant from an existing prototype, the perceptual space is then entirely taken up by the existing categories. Hence, a new category has *no room* in the perceptual space, which is thus being perceived as a token of any of the already existing categories.

2.4.2.1 A psycholinguistic account: The Automatic Selective Perception model (ASP)

The Automatic Selective Perception model (ASP) (Strange, 2011) is an empirically-based account of the differences in the processing of native and non-native sounds. The ASP model proposes that listeners employ selective perception routines (SPR), which aim to detect the relevant information in the L1. These perception routines are developed during the L1 acquisition process and become automatic throughout time. The ASP also defines two different modes of perception: phonological (which is mostly used in the L1) and phonetic. The phonological mode "enables the listener to detect *sufficient* phonologically relevant contrastive information for word-form identification" (p. 460). As mentioned above, this mode of perception is over-learned, automatic, and allows for quick, robust lexical access through specific perceptual attunement to the relevant auditory cues.

On the other hand, the phonetic mode of perception (where listeners can perceive acoustic detail that the phonological mode does not take into account) is less frequent and is more cognitively demanding than phonological mode insofar as it requires a certain degree of 'undoing' the perceptual routines that lead to optimal perception in the L1. In this sense, the ASP claims that L2 listeners still make large use of their L1 perceptual routines, but that certain perceptual tasks and conditions could trigger a phonetic mode of perception that would in turn increase performance in the L2.

This model assumes two different modes of perception which chiefly correspond to the usual phonetics/phonology distinction in production and justifies the phonetics/phonology distinction in *perception* that this work relies on. Phonological mode is then comparable to a categorisation task, in which listeners make use of a perceptual routine specifically tailored to interpret the input as members of *existing* categories. Phonetic mode, on the other hand, does not necessarily make use of the existing categories as it allows for perception and processing of otherwise unnoticed sub-categorial detail.

If we assume that phonological mode is the default perceptual mode, then creation of nonnative categories would require a change of perception mode that naturalistic input does not trigger. I therefore claim that creation of new phonemic categories via noticing requires a phonetic mode of perception, which would make this process possible. However, phonetic mode is not likely to occur in regular situations of communication; it needs to be *triggered* by specific perceptual tasks such as discrimination.

Nevertheless, and as I mention in section 2.5, this phonetic mode of perception may still not be a sufficient condition to make category creation possible, because there could be no reason for L2 listeners to change their perceptual routine. Therefore, they will not have access to processing the acoustic detail that will lead them to create new categories.

2.4.2.2 Categorical perception

Categorical Perception (CP), in its original version, refers to the relation found between an identification and a discrimination task, where identification predicts discrimination. Iverson and Kuhl (2000) point out that CP has also been understood as a perceptual mechanism:

> The term *categorical perception* is mostly used as an empirical description of the correlations between discrimination sensitivity and phoneme labelling in perceptual experiments, but it is strongly linked with a class of hypothesized mechanisms in which phonemes are per-

> ceived in terms of their phonemic labels rather than by their acoustic properties. (p. 874)

Categorical perception tests can be traced back until the late 1950s, with the seminal work of Liberman et al. (1957) and it refers to the effect found in the /b-d-g/ continuum, where listeners would discriminate more accurately between adjacent tokens that were on different sides of a boundary previously defined by a labelling task.

In their experiment, Liberman et al. used synthesised stimuli, creating a 14-step continuum along the voiced stop + vowel sequence /ba/, /da/, and /ga/, with the second formant (F2) as the varying perceptual cue. Two tasks were performed. In the identification task, subjects were asked to categorise each one of the stimuli into the consonants *b, d, g.* The second was an ABX discrimination task, with stimuli separated by an Interstimulus Interval (ISI) of 1 second[3]. Discrimination was made between immediately adjacent stimuli (1-step discrimination), and also between stimuli 2 and 3 steps away. The results showed that subjects identified stimuli along the continuum *categorically*, that is, that categories had clear, sharp boundaries along the continuum. According to the results, the boundary between categories /b/ and /d/ was at token 4, and the boundary between /d/ and /g/ between tokens 9 and 10[4]. On the other hand, the discrimination task also showed a consistent pattern: Discrimination raised significantly in the places where the identification task had set the category boundaries.

The main claim of this work was that discrimination between two stimuli was better when each stimulus corresponded to instances of two different categories, and that within-category stimuli were often perceived as being the same. Therefore, discrimination was determined by categorisation, and as such, one should not expect that subjects could be able to discriminate between sounds that sit within the same category. However, Liberman et al. do not consider that this should be held as true in all cases, and neither do scholars that have done research on this effect later on. Further replications of this experiment showed that vowel perception would not show the same pattern (Fry, Abramson, Eimas, & Liberman, 1962), and that the type of discrimination task would also play a role: The more bias-free the task is, the less CP effects are shown (Gerrits & Schouten, 2004).

As Iverson pointed out, this experiment led to two different concepts of CP. On the one hand, a CP experiment may refer to one where identification and discrimination tasks along a continuum are carried out, and on the other hand, it refers to the underlying *perceptual mechanism* that causes listeners to perceive sounds according

[3]The Interstimulus Interval – the period of time between two given stimuli within a trial – is used in discrimination tasks due to its supposed importance to the type of knowledge accessed; see section 3.4.1.2.

[4]By ‘boundary’ we mean the point in the perceptual space where the probability to perceive a given token as a member of one out of two adjacent categories is 50%.

to the existent phonemic categories, hence reinforcing the concept of perception of linguistic sounds as a categorisation task. The latter definition of CP is crucial, as it predicts discrimination behaviour (which is label-free) to depend on the phonemic labels. This view poses a problem with regard to category creation in the L2, as this needs to happen in a reverse direction: In order to create new categories, L2 learners need to stop relying on the phonemic categories of the L1 and start perceiving phonetic detail so that the noticing of the differences between their perceptual representations and the actual input takes place.

However, an important question arises: is CP the same for vowels and consonants? In this regard, experiments on categorical perception of speech sounds have shown that while consonant discrimination triggers a sharp boundary effect in native speakers, vowels do not seem to follow this pattern, possible due to the number of perceptual cues being manipulated, or the level of dispersion found in formant values of vowels (Fry et al., 1962).[5] The lack of sharper boundaries in identification and higher discrimination along the entire continuum obtained by Fry et al. has been considered as evidence that vowel perception is not categorical, a claim that has been confirmed by Schouten, Gerrits, and van Hessen (2003, p. 79). But the lack of categorical perception does not mean that there are no categories. Rather, it means that the *nature* of these categories is not the same: categories may also have diffuse boundaries and overlapping.

Furthermore, Fry et al. raise an interesting hypothesis about perception of English vowels: given the high level of dispersion of their formant values in the acoustic space due to dialectal differences, it might be the case that speakers of this language are specifically trained to “shift the frame of reference for vowels very frequently and very rapidly under the influence of sounds heard in the immediate past” (Fry et al., 1962, p. 188). In this regard it is worth noting that the acoustic values for /a/ were the same as those for /ʌ/ in Experiment 1, which is also seen in the data by Bradlow (1995).

2.4.2.3 The Perceptual Magnet Effect (PME)

A different approach to speech perception in L2 is that of a prototype-theoretical framework, where perceptual categories are defined by *prototypes* (best exemplars) of the sound in question. The concept of prototype in cognitive psychology was coined by Rosch (1999), who defines it as “good instances” of a certain category.

[5] It is perhaps worth noting that the effects of the *number* of perceptual cues available is different if the *nature* of these cues is different: a continuum built between two vowels such as /i/ and /ɪ/ where duration and F1 are being modified should not necessarily yield the same results in discrimination or identification as a continuum such as /ɑ/ and /ʌ/, where F1 and F2 are both spectral cues. However, the availability of duration as a perceptual cue in the language of the speaker does not necessarily have to influence the results: the work by Escudero (2005) suggests that L2 learners of English with Spanish as L1 could make use of the duration cue as a bootstrapping mechanism.

Prototypes are relevant in cognitive accounts of categorisation as they have a unique perceptual status: They are easier to classify, easier to remember, and preferred over the other members of a category (Kuhl & Iverson, 1995, p. 123).

Research on the Perceptual Magnet Effect (Kuhl & Iverson, 1995) brings this prototype-based approach to the field of speech perception and proves that perceptual categories that are linked to phonemic categories in the native language also display this prototype effect.[6] According to previous findings by Kuhl (1991) on phonetic perception in both adults and infants, perception of vowel sounds is altered (or 'warped') as a function of exposure to language during the earliest stages of first language acquisition (6 months).

The PME consists of a lack of discrimination ability between tokens of the same vowel, when one of these tokens is considered to be a prototype (or 'good exemplar') of that vowel, and the second token is very close to it in the acoustic space. When subjects are asked to discriminate between these two sounds, they perceive it as the same. Conversely, when a non-prototypical token of the same vowel is paired with a token that is equally close in the acoustic space, the results are different; subjects in this case are able to perceive these two sounds as different. Hence, prototypes of vowels work as 'perceptual magnets' in a way that non-prototype sounds do not.

The experiment carried out by Kuhl (1991) included tokens from a synthetic continuum between a good examplar and a bad exemplar of /i/, and subjects were asked to discriminate between either the prototype and its surrounding variants, or the nonprototype and its surrounding tokens. While this experiment was performed in adults, infants, and rhesus monkeys as well, infants were tested using the head-turn preference procedure (where they were reinforced in their response by showing them a toy); adults were asked to press a button. Monkeys had to operate a response key to show that they had noticed a difference, and were recompensed with food instead of a toy. The reason why rhesus monkeys were chosen was because they share certain similarities in perception, such as the categorical perception phenomenon (cf. section 2.4.2.2 above). However, the specificity of the Perceptual Magnet Effect resulted to be higher than CP, as monkeys did not show any traces of perceptual warping during the test.

These findings led Kuhl and Iverson (1995) to claim that exemplars of phonemic categories act as perceptual magnets, for which the stimuli surrounding the prototype would be perceived as members of the same category; conversely, poor exemplars of the category do not "attract" the surrounding sounds in the continuum. This effect has important implications for L2 speech perception, which were confirmed by a different experiment by Kuhl et al. (1992). In this research Kuhl et al.

[6]It is worth noting that the PME is tested on vowels that have phonemic status in native speakers of a language, which is not the case of phonetic categories that belong to the same phonemic representation and are in complementary distribution (i.e. in categories such as [p^{h}] and [p]).

tested the perception of tokens of American English /i/ and Swedish /y/. Subjects were 6-month-old infants raised monolingually in Sweden and the US. American infants showed a strong perceptual magnet effect for /i/, with their surrounding tokens being perceived as being the same; conversely, surrounding tokens of /y/ were discriminated as different, given that the /y/ sound is not part of the linguistic experience of a monolingually raised American child. Swedish infants on the other hand showed the opposite pattern, with a strong perceptual magnet for /y/ and no effect for /i/.

A further study by Iverson et al. (2002) showed that nonnative sounds that are categorised as members of the same native category will be particularly difficult to perceive as different, as they are being attracted by the prototype of the native category. This experiment tested perception of an /r/ - /l/ continuum that varied in F1, F2, and F3. Subjects were a group of Japanese speakers, a second group of American speakers, and a third group of German speakers, all of them raised monolingually. All participants had to identify the tokens as members of their native phonemic categories, and rate them on a scale from 1 to 7. Later, subjects rated pairs of tokens in terms of their similarity, on a scale from 1 to 7. Results showed that Japanese listeners classified their tokens as members of one L1 category, /l/, though they were able to discriminate in a within-category manner. American listeners, on the other hand, showed a high perceptual magnet effect for each one of the category prototypes. Finally, German listeners showed a similar effect than that of American listeners, with the exception that Germans classified tokens of /r/ as exemplars of the German uvular fricative.

In sum, these two approaches have shown that speech perception itself offers difficulties in terms of creating new perceptual categories, since listeners rely on their phonemic categories to process speech sounds. If, as the CP framework suggests, discrimination is weaker for within-category tokens then L2 speakers would not be able to perceive the acoustic differences and would therefore continue mapping the incoming input as members of their native category. Likewise, since the PME suggests that discrimination is difficult when comparing tokens of the same category as they would be perceptually warped onto the L2 speaker's native category, then the probability of creating new perceptual representations on the basis of incoming acoustic input only seems to be rather low. Furthermore, if on top of these perception-specific constraints we add the maturational and feedback-related constraints, then the creation of a new perceptual category seems to be one of the most challenging tasks in L2 learning.

2.5 Case study: the /ɑ/ - /ʌ/ contrast in native speakers of Spanish

So far I have discussed the reasons for which acquiring a new perceptual category is one of the most difficult tasks when learning an L2, and I have grouped these constraints into three main categories: maturational, psychoacoustic, and input-related. Here, and in order to understand the potential development of new perceptual categories, I present a specific case study and describe it from a phonetic and phonological perspective.

2.5.1 Spanish and English vowel systems

Spanish has a very simple vowel inventory, compared to English: While Spanish has only 5 monophthong vowels in all of its dialects (Quilis, 1993, p. 144), English shows a much more complex system that can range between 10 and 13 monophthongal vowels according to the dialect (Giegerich, 1992). The present study takes as reference a relatively small English vowel system: A General American English (GAE) dialect with a COT/CAUGHT merger, i.e. a system that does not make a phonemic distinction between a open, back, unrounded vowel /ɑ/ and a near-open, back, rounded vowel /ɔ/.

In the perceptual/articulatory vowel space the Spanish system takes the shape of an inverted triangle, with a three-way height (openness) distinction: high (close), mid, and low (open). The second distinction is frontness, with front, central and back values. Rounding is redundant and only present in back vowels /o/ and /u/. Table 2.1 summarizes the phonetic labels of the Spanish vowel system.

On the other hand, the American English vowel system takes a trapezoidal shape in the articulatory space, with a two-way frontness distinction (front and back, though with one mid, central reduced vowel /ə/), and a four-way openness distinction (close, close-mid, open-mid, and open). Phonologically speaking, vowels /ɪ/ and /ʊ/ are considered to have the same features as /i/ and /u/ with the addition of a lax feature, a distinction that the IPA refers to as nearness and thus presents the tense/lax distinction as a different degree of openness and frontness. Table 2.2 presents the articulatory labels for monophthongal English vowels in a GAE system; Figure 2.1 shows both vowel systems plotted one next to each other.

The /ɑ/ - /ʌ/ distinction poses two challenges for a learner with an L1-Spanish perceptual grammar. Spanish does not have a near-open/open height distinction as English does, for which non-native vowels that fall in the near-open area will have to be perceived as either a mid vowel or the open vowel. At the same time, the only back vowels in Spanish are /o/ (mid) and /u/ (close), with /a/ having a central value in terms of frontness. Hence, both the open, back vowel /ɑ/ and the

Table 2.1: Spanish vowel system.

Vowel	Height	Position	Roundness
i	close	front	unrounded
e	mid	front	unrounded
a	close	central	unrounded
o	mid	back	rounded
u	close	back	rounded

Table 2.2: American English vowel system (no COT/CAUGHT merger).

Vowel	Height	Position	Roundness
i	close	front	unrounded
ɪ	near-close	near-front	unrounded
e	close-mid	front	unrounded
ɛ	open-mid	front	unrounded
æ	near-open	front	unrounded
ɑ	open	back	unrounded
ʌ	open-mid	back	unrounded
ɔ	open-mid	back	rounded
ʊ	near-close	near-back	rounded
o	close-mid	back	rounded
u	close	back	rounded

near-open, back vowel /ʌ/ could be perceived as either an open central vowel /a/ or a mid back vowel /o/.

However, discovering which L2 vowel is mapped onto which L1 vowel is not simple and the possibilities should not be taken for granted. An important problem is determining whether an L2 vowel that perceptually lies exactly on the boundary between two L1 categories A and B, will be mapped onto A or B. If one takes into account the assumption that the answer is a matter of probability, then a certain universe of subjects would show a 50/50 response pattern. But it might be the

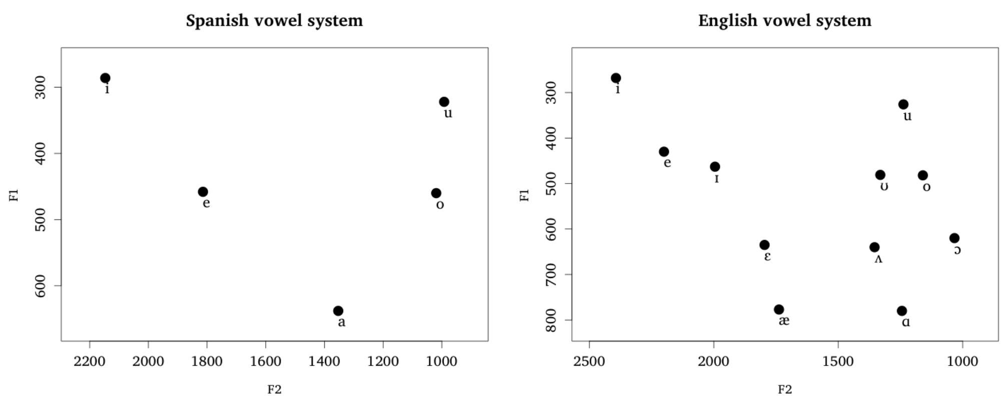


Figure 2.1: Spanish and American English vowel systems (data from Bradlow, 1995).

case that speakers might have become more sensitive to any of the perceptual cues involved in the perception of the L2 vowel, for which it would be expectable to find a shift from vowel A at the initial stage, to vowel B at a more advanced learning stage. In addition, this choice might also be subject to dialectal variation: If different dialects of the L1 have different cue-weighting strategies, then different dialects of the L1 will have different mappings. Furthermore, a third cause for different mappings can be considered if L2 speakers are able to listen to different variations of the TL. This work attempts to control for these variables by considering speakers of one L1 dialect only, who received input from mostly one variety of the TL. The way in which these possible initial states evolve into IL representations will be discussed in Chapter 3.

From all these options in terms of the initial state, I will assume that the perceptual mapping of the English vowels /ɑ/ and /ʌ/ is such that both are perceived as the L1 vowel /a/, in what PAM would call Single-Category assimilation, or as I will also call it here, a *2-to-1 mapping* scenario[7]. Further experimental support for this initial state hypothesis can be found in Escudero and Chládková (2010) and García (2014).

I will assume the approach proposed by Escudero (2005) in that optimal L2 speech perception is the way in which a native speaker of English will perceive this pair of sounds, namely, as: (a) two different phonemic categories; and that (b) at least one of these categories has target acoustic values and/or articulatory features that do not correspond to any of the categories that are present in Spanish (p.89).

This case is a good example of sub-optimal L2 speech perception where all the constraints mentioned throughout this chapter apply: (a) a learning situation involving late-learners; (b) two L2 sounds being perceptually mapped onto one L1 category; and (c) no previous phonetic training. While vowels /ɑ/ and /ʌ/ are contrastive due to their difference in height and duration (from a phonetic perspective), and tenseness (from a phonological perspective), native speakers of Spanish have no back unrounded vowels, no open-mid vowels, nor tense/lax contrasts, which causes these two vowels to be perceived as either a low vowel /a/ or a mid back rounded vowel /o/. Chapters 4 and 5, however, provide evidence that both vowels are perceived as /a/.

2.5.2 A featural approach to perception

While most studies in L2 speech perception adopt an acoustic similarity approach, fewer have considered a featural approach, that is, understanding perception

[7]NB. When I refer to '2-to-1 mapping' I am describing the English-to-Spanish category mapping, and more specifically, that these two English vowels are being mapped to one Spanish vowel. This is different than 'perceptual mapping' or 'acoustic-to-linguistic mapping', which refers to the mapping of the actual acoustic signal (i.e. formant frequencies in the case of vowels) onto the linguistic categories that are present in the L2 learner's IL, whatever these are.

of L2 sounds according to the phonological features present/absent in the L1/L2; in this regard, it is worth considering the work by Brown (1998), where she claims that feature contrasts in the L1 can be redeployed to the L2, thus ensuring target-like perception. However, the predictions for L2 speakers whose L1 did not have the features needed to correctly perceive a certain L2 contrast were more pessimistic as it was assumed that the L1 *interferes* with UG access.

Such interference of the L1 in acquisition via UG applies to this case study. The fact that the Spanish vowel system does not include the [tense] feature makes acquisition impossible under the tenets of Brown's work, given that phonologically, the relevant vowel features involved in perception of the /ɑ/ - /ʌ/ distinction are [high], [low], [back], and [tense]. On the other hand, and despite the simplicity of the inventory, the feature values for the Spanish vowel system are somewhat less clear. The reason for such lack of clarity can be found in the features corresponding to the vowel /a/, precisely the category onto which the vowels /ɑ/ and /ʌ/ are being mapped. /a/ could then be interpreted in three different manners in terms of frontness/backness: (a) as a [+low][+back] vowel; (b) as a [+low][-back][-front] vowel; or (c) [+low] and simply underspecified for frontness/backness. In this regard, Núñez Cedeño and Morales-Font (1999) point out that the Spanish 5-vowel system does not need a [front] feature, since it is usually taken for granted that /a/ is [+back]; however, neither production nor perception are constrained by this hypothetical [+back] feature of /a/: while [æ] is still possible in production, Spanish speakers also perceive the sound as an exemplar of /a/[8]. Hence, the only reason for this specification relies on the theoretical need for binarity (p. 38-39). Table 2.3 shows the values for each one of these English vowels as presented by Giegerich (1992), and those for Spanish vowels as presented by Núñez Cedeño and Morales-Font (1999).

Table 2.3: Features involved in the /ɑ/ - /ʌ/ distinction in English, and the /a/ - /o/ distinction in Spanish.

	/ɑ/	/ʌ/	/a/	/o/
[High]	−	−	−	−
[Low]	+	−	+	−
[Back]	+	+	+	+
[Tense]	+	−	0	0
[Round]	−	−	−	+

I have also presented Spanish /o/ as a category onto which either one of the English vowels /ɑ/ and /ʌ/ may be mapped. The features for this vowels are [+back], [-high] and [-low], and [+round], where rounding offers a third perceptual cue to the listener (i.e. F3). Hence, from a feature similarity viewpoint, /a/ is a

[8]Furthermore, Navarro Tomás (1974) noted the presence of a retracted low vowel [ɑ] and a raised [æ] in Spanish (p.55-56). This has also been empirically observed by Martínez Celdrán and Fernández Planas (2007), who concluded that their values are within the dispersion field of /a/ (p.183-185).

good candidate to map /ɑ/ onto insofar as it is also [+low] and [-round], although (maybe) not [+back]; and on the other hand, /o/ is the best candidate to map /ʌ/, given that they share the [-low] and [+back] feature, although not [+round].

If the candidates are reversed then /ɑ/ could also be mapped onto /o/, given that they share the [+back] feature, but /ʌ/ is less likely to be mapped onto /a/ given that they only share [-round] and a potential [-front] feature, which is shared by all the vowels in question.

It seems then, that a feature-based approach to perception does not predict a Single-Category assimilation, but a Two-Category one. As for learning, the L2 speaker of English with Spanish as L1 would have two options in order to acquire the contrast: (a) to either add a [tense] feature to their phonological knowledge, or (b) to add a new member by copying the features of /o/ and changing its [+round] feature to [-round]. However, would Spanish speakers be able to perceive this [-low] feature without [+round] being also present? Furthermore, GAE has a different articulatory target for /ʌ/ than British English, thus acquiring a less back phonetic realisation, closer to [ə]; this degree of frontness makes it again closer to Spanish /a/, which would then explain the Single-Category assimilation scenario.

Hence, feature- and acoustic-based approaches have different predictions for acquisition of L2 vowels, which leads to the question: which approach is more explanatory? A study by Barrios et al. (2016) on perception of /i/ - /ɪ/ and /æ/ - /ɑ/ contrasts by native speakers of Spanish showed that neither an acoustic nor a phonological similarity approach to perception seemed to be necessary nor sufficient in order to explain the perceptual patterns of the subjects. Barrios et al. noticed that while both vowels of the /i/ - /ɪ/ distinction are perceptually assimilated by L1 Spanish speakers of English as /i/, their F1 values show that /ɪ/ is acoustically closer to the Spanish category /e/. Hence, the acoustic approach does not make the right predictions. In the following chapters I will provide evidence that even though /ʌ/ is acoustically closer to /a/ (for which it is expected that /ʌ/ will be considered a better exemplar of /a/ than /ɑ/), nonnative speakers are almost equally likely to perceive tokens of /ʌ/ as exemplars of /o/, whilst tokens of /ɑ/ (whose formant values are farther from /a/ than those of /ʌ/) are unequivocally categorised as exemplars of /a/.

2.5.2.1 The learnability issue of category splitting

The 2-to-1 mapping scenario poses several problems in terms of learnability. First of all, it requires a *split* of two perceptual L1 categories, which is one of the most complex learning scenarios. A split, according to Escudero (2005), is an operation that requires two types of learning: a perceptual type of learning that requires a modification of the perceptual mappings of the L1; and a representational one, entailing the creation a new category (p.158).

The fact that perception is warped in a non-linear way makes perceptual strategies more difficult to learn. In an ideal world where humans perceive linearly, sounds would be perceived with great phonetic detail. But we know that in the real world, perceptual categories have a unitary identity that makes within-category phonetic detail irrelevant for purposes of lexical and phonemic identification[9]. Whenever the continuum of the perceptual space is segmented into discrete perceptual categories, the specific acoustic values of the stimuli become "blurred", or in other words, they assimilate to the prototypical values of the category, as explained by the NLM.

However, category splits cannot be explained in a one-size-fits-all manner. The learnability of the split will depend firstly on the perceptual cues that are available in the L1, as well as on the potential learning of a different, L2-specific perceptual cue. Secondly, perceptual cues have a different saliency, for which having more perceptual cues does not necessarily result in easier learning; in this regard, duration might be a universally salient perceptual cue (Bohn, 1995, as cited in Escudero, Benders, & Lipski, 2009), which can be used by nonnative speakers as a way to create a category split where the target categories are different in both duration and vowel quality (Escudero, 2005). I therefore predict that when the cues that lead the distinction are (a) also present in the L1, but at the same time (b) do not lead to any feature distinction in the L1, then learning reaches the highest degree of difficulty. This prediction applies to the /ɑ/ - /ʌ/ contrast, in which one of the relevant acoustic cues is F1. This cue is also present in Spanish but does not lead to discrimination in terms of openness: the values of both /ɑ/ and /ʌ/ are *within* the range of values that corresponds to the only low vowel in Spanish /a/. Furthermore, the second available cue, F2 (related to frontness/backness) is also present in Spanish, but again, the F2 difference between /ɑ/ and /ʌ/ does not lead to distinctions in a 5-vowel system as the acoustically closest front/back contrast (i.e. between /e/ and /o/) is given by differences of over 1500 Hz. Finally, while duration is also available in the input given that /ɑ/ is longer than in /ʌ/ in GAE (J. Hillenbrand, Getty, Clark, & Wheeler, 1995), Spanish does not have a duration contrast in either vowels or consonants[10]. A second consideration regarding duration is that according to Strange et al. (1998), vowel length in American English is considered a phonologically redundant feature which varies according to the phonological context (p. 316). This claim is supported by perceptual tests showing that native speakers do not rely on it as a primary source

[9]It is worth noting that this statement does not generalise to L1 perception, particularly in sociophonetics, where speakers are able to perceive subphonemic detail. This particular case of category splitting in L2 assumes that the listener has not been exposed to an input that consistently shows a certain subphonemic n-modal distribution pattern, as it is the case of language variation in L1.

[10]The fact that the Spanish phonological system does not have phonemic length contrasts nor uses duration as a cue is not a trivial point, since it has been shown that L2 speakers who have a duration distinction in consonants but not vowels in their L1 may be able to redeploy this feature to an L2 with vowel length distinctions. I present a few examples of feature redeployment in section 3.3.1.2.

for vowel identification (see Karpinska, Uchida, & Grenon, 2015, for cross-dialectal comparison of the *bit* - *beat* contrast, and J. M. Hillenbrand, Clark, & Houde, 2000, for vowel identification by phonetically trained listeners). Given the high variability of duration in vowels, it is unlikely that nonnative speakers would consider duration as a relevant cue for the purposes of acquiring a category contrast. In this regard, it is worth mentioning the study by Escudero and Boersma (2004), which shows that experienced L2 speakers move towards target-like weighting of the acoustic cues; that is, they tend to copy the perceptual cue-weighting of the native speakers of the TL.

This fact leads to a bootstrapping situation, where the L1-optimal perceptual strategies do not allow L2 listeners to perceive the perceptual cues involved in the relevant vowel distinction: the /ɑ/ - /ʌ/ distinction in the TL input is then not perceived as two different sounds. Hence, and in the initial state, lexical access to minimal pairs is only cued by semantic context, for which the need to learn the perceptual contrast becomes dependent on the number of minimal pairs in the L2 and their likelihood to appear in a similar semantic and syntactic context.

We have previously mentioned that perception of speech sounds in phonological mode is ultimately a categorisation process specifically designed to deal with variation and noise in a rather robust manner. Since this is the case, then perception of differences in acoustic detail is not a task that the default speech perception mechanism is good at: Input has to be categorised regardless of the potential deviance of the signal. Perceiving within-category differences is in fact a process that goes *against* the natural drive to categorise the sounds in order to obtain meaning, and thus it is cognitively discouraged as a default. But this is exactly what an L2 speaker who attempts to create a new perceptual category needs to do in order to avoid the automatic drive to categorise the acoustic input into the existing L1 categories: to tune in the ear for fine acoustic detail that will help her to notice the relevant acoustic detail.

If we are to assume that the learner took mental note of this specific token and its deviance from a given L1 sound prototype, then we could expect that this process would repeat until the amount of deviant tokens noticed by the learner is large enough so as to lead her to think that these deviant tokens are not just a product of chance but are actually a different category in the L2. Nevertheless, in order to reach this stage the first step (that is, going against the default categorisation setting) needs to be taken; in this regard, the notion of *noticing* seems to be crucial for this process to take place.

Furthermore, the notion of *intake* (defined in section 2.2.3.1) in L2 does not seem to apply to the domain of phonology in the same way as in syntax. While noticing seems to be an important element to bridge the gap between the received input and the intake, there are important limitations in the concept that are exclusive to the sound level of language: While some specific L2 acquisition strategies such as input enhancement work quite well in certain domains such as syntax and morphology, the phonetics/phonology level is not one of these. As it will be shown in the following section, perception of acoustic stimuli and their categorisation into speech sounds is mediated by a perceptual grammar, which is (a) language-specific, and (b) needs to deal with great variability in the stimuli by operating constraints that force categorisation into the existing phonemic categories.

Hence, non-native perception of speech sounds will have to inevitably undergo a sub-optimal speech perception stage, which is characterised as a full transfer of the phonological system of the L1 to the blank slate of the L2. This does not only imply transfer of the phonemic inventory; it also includes the *perceptual strategies* involved in optimal L1 speech perception. In order to understand this notion we will need to consider the concept of a perception grammar.

2.5.2.2 The perception grammar and its role in L2 speech perception

We will follow Escudero's (2005) view of speech perception in L2 as a twofold task, consisting firstly of the creation of perceptual categories, a process that takes place during the acquisition of a language, be it the L1 or an L2. And secondly, it entails a perceptual mapping of the acoustic signal onto these linguistic categories. This model, as we will see in Chapter 3 in more detail, is based on the tenets of Functional Phonology (Boersma, 1998). Boersma defines a perception grammar on the basis of an Optimality-Theory framework (Prince & Smolensky, 2004), where the selection of potential candidates in production is given by a set of universal constraints which are ranked differently according to the language. Boersma adopts this approach for perception, thus presenting it as a process where categorisation is a result of interactions among ranked constraints that relate to acoustic values and the way in which these are mapped onto categories. In order to understand the changes in perceptual mappings and representations that allow for a target-like perception in L2, the concept of perception grammar becomes crucial:

> The perception grammar classifies continuous auditory inputs into a finite number of discrete categories, by means of an interaction between several constraint families (Boersma 1998: 161-172, 336-343, 375-379). Each candidate is evaluated for its conformance to the language-specific set of phonological feature values (*CATEG) and for its closeness to the acoustic event (*WARP). (Boersma, 1999)

Hence, a perception grammar has:

- An **acoustic input**, corresponding to the acoustic signal received by the listener (e.g. formant values);

- A **set of candidates**, i.e. the perceptual categories available in the listener's phonological inventory; and
- A **set of perceptual constraints**, which operate over the properties of the acoustic input and regulate the mapping of the acoustic input onto the corresponding category.

However, Boersma's model is esentially an L1 category acquisition model, and as such it is not concerned with creation of categories in adults (that is, L2 acquisition). More importantly, the constraints that correspond to this model (mentioned above) do not refer to abstract categories, for which Escudero (2005, p.46) defines them as *auditory-to-auditory* constraints.

The model was then extended by Escudero and Boersma (2003), which was later known as the Linguistic Perception model (LP) (Escudero, 2005), where speech perception is defined as an auditory-to-linguistic process. In Escudero's work, perception of speech sounds is modelled as a mapping process from raw acoustic input to phonological categories at the segment level. This acoustic-to-phonological integration is mediated by *cue constraints*, which "explain how multiple auditory cues are perceptually integrated in order to map the speech signal onto phonological categories" (Escudero, 2009, p. 155). The constraints in the LP model then look like 'X Hz is not /category/', which stress the auditory-to-phonological nature of language-specific constraints.

The L2LP model (Escudero, 2005) is a further elaboration of LP, but applied to L2 speech perception. The set of constraints posited by this model are the same as those in LP. Here, L2 speech perception is modelled as a process where the L2 learner moves from a full L1 copying state to optimal L2 speech perception (p. 89), as suggested by Schwartz and Sprouse (1996). However, this type of constraint does not make explicit the way in which new perceptual candidates may appear. This problem could be solved by taking the *Categ from Boersma's first model and apply it to L2 speech perception. In this regard, the assumption of a high-ranked *Categ constraint family in adults would militate against the creation of new perceptual categories whenever a listener receives input, but at the same time, if *Categ is violated, then a new category could be created. Likewise, L2LP does not explore the possibility of using *Warp constraints, which as it will be shown in Chapter 3, could allow for modelling subcategorial learning.

2.6 A Single-Category assimilation model

In a model of L1 acquisition, Boersma (1997) proposes category creation as a result of distributional learning. Categorisation is governed by constraints that take into account not only the possible acoustic values corresponding to the existing categories, but also their frequency in the input. However, the set of constraints in this model is ranked in such a way that PERCEIVE is the highest ranked constraint, for which anything recognised as a speech sound will be more likely to be categorised sub-optimally into whatever existing category there be, rather than being non-categorised; furthermore, the *CATEG constraint family ensures that acoustic input will be categorised into *existing* phonological categories.

On the other hand, the L2LP approach uses only acoustic-to-phonological constraints, which makes acquisition of new phonological categories problematic: why are new candidates being created? If, as it has been assumed by both models, creation of categories is given by naturalistic input only, then something else is needed in order to explain the generation of more candidates by GEN.

Figure 2.2 shows the role of the perception grammar in L2 speech perception: the input is taken by the perception grammar, which categorises the sounds according to the existing perceptual constraints and their ranking. This representation of L2 speech perception assumes an initial state where the L2 learner has no awareness that the TL vowel inventory is larger than that of the L1. When the listener perceives the incoming input, the acoustic values of it are received by the perception grammar, which is set with the parameters of the L1, both in terms of the mapping of the incoming acoustic values and the available perceptual representations.

However, the output of the perception grammar (i.e. categorisation) can take many forms, thus leading to different perceptual outcomes. In this regard, I propose four possible states of the perception grammar, which will lead to four different perceptual outcomes. The first stage is given by a full copying of the L1 perceptual grammar, which will evaluate the acoustic values of the incoming input and categorise them into the existing L1 categories; nevertheless, they will be considered poor exemplars of them. I predict that such a stage would lead listeners to perceive both /ɑ/ and /ʌ/ as members of the Spanish category /a/; this is what causes the 2-to-1 mapping situation. A second state of the perception grammar allows the listener to identify the L2 sounds as exemplars of different L1 categories; this is, no new categories have been created, but they perceive the incoming input as members of a different neighbouring L1 category. In this case study, this would cause the listeners to start remapping one of the L2 categories onto a different L1 category, thus remapping, for instance, /ɑ/ onto /o/, and leave /ʌ/ as the L2 counterpart of /a/. The third state gives way to creating subsets of existing L1 categories; that is, listeners would start noticing that the L2 has ’two kinds of /a/’. Finally, a fourth state would imply one further step from the previous one: the grammar will eventually include a new category as a candidate for categorisation of the incoming input, where one of these ’two kinds of /a/’ would become a category of its own.

In this regard, it is worth noting that the exact way in which the perception grammar creates a category according to the existing models is unclear. While the Gradual Learning Algorithm (GLA) (Boersma, 1997; Boersma & Hayes, 2001) was originally formulated for first language acquisition and is mostly concerned about the way in which speakers/listeners rerank the existing constraints, Escudero (2005) extends the validity of the model for acquisition of L2 categories. Nevertheless, the GLA requires the infant to rank the *Categ constraints very low, in order to make category creation possible. However, I have also claimed above that presumably adults have ranked this constraint very high; in this regard, the point at stake is whether adults can demote the *Categ constraints so that their L2 category acquisition process becomes possible; Chapter 3 will analyse the possible set of contraints involved in category acquisitions in cases where two TL vowels are mapped onto the same L1 vowel.

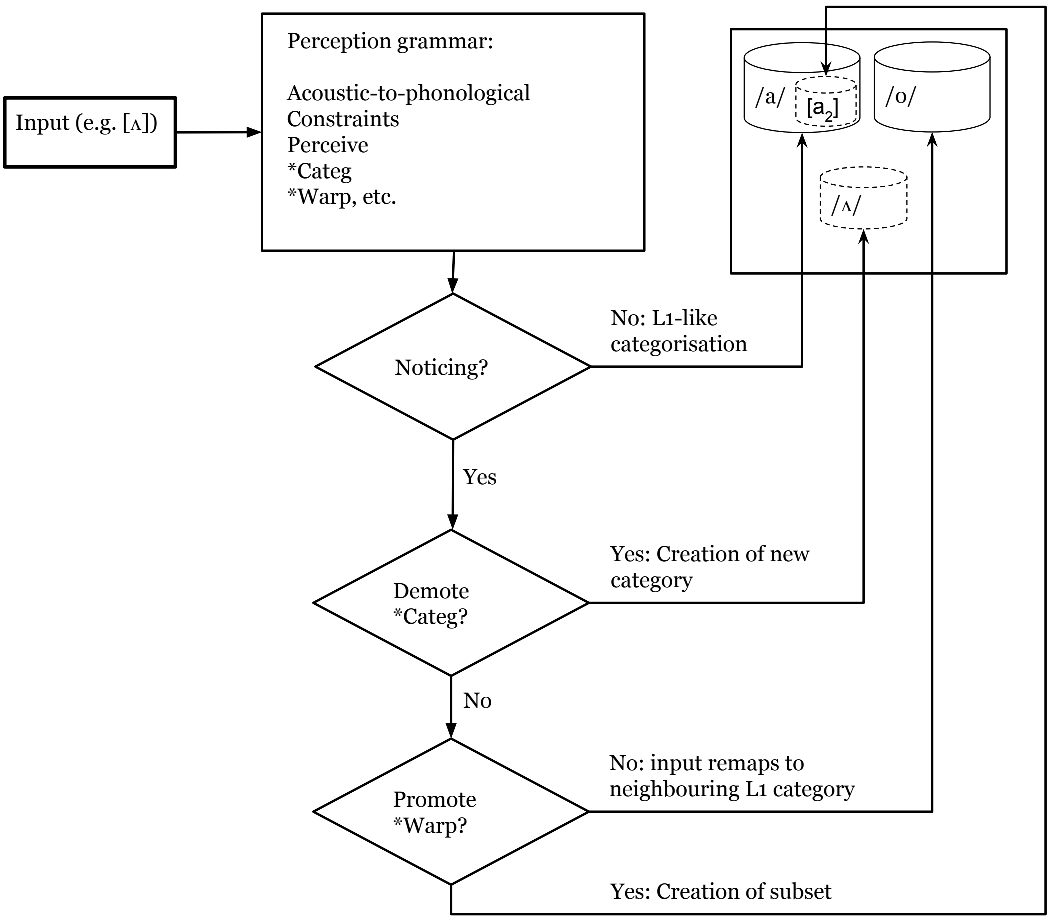


Figure 2.2: The perception grammar in L2 categorisation and category creation. Note how this formalisation still cannot properly explain what exactly triggers creation of new categories. Here it is assumed that noticing *may* lead to category creation, but the mechanism that allows it in favour of less optimal outcomes is unclear.

Therefore, the copied perception grammar can follow four paths. However, it is only after these acoustic constraints operate that the cue constraints proposed by L2LP can be invoked and thus reach the categorisation goals.

(a) To ignore the acoustic detail in the incoming L2 input by means of a very low ranked *Warp constraint, thus assuming L2 input as members of the L1 categories;

(b) To notice such acoustic differences and adjust itself in such a way that no new categories are created, but with some acoustic values being mapped onto other existing categories;

(c) To notice the acoustic differences, but still encode the perceived differences in the grammar as a subcategory conformed by less prototypical exemplars of the L1 category, due to a highly ranked *Categ constraint; or

(d) To notice the differences and demote *Categ constraints that allow for the creation of new categories.

Option (a) will lead to sub-optimal L2 speech perception, which is assumed to be the initial state of perceptual category acquisition. The incoming tokens of both /ʌ/ and /ɑ/ will be warped by one of the native categories (presumably the one with the closest target values in the acoustic space) and will therefore be categorised as such. If we make the assumption that this mechanism repeats in time, it will eventually lead to a state of fossilization. It can be hypothesized that the decision made by the perception grammar in terms of categorisations feedbacks onto itself, thus discouraging further changes: if the grammar decided that /ʌ/ is a valid exemplar of /a/ and the communication goals are met, then there would be no need to change this mapping.

On the other hand, the options where a noticing process has taken place could lead to optimal L2 speech perception, which for the sake of exposition will be understood as native-like perception (more details on this concept are provided in Chapter 3). However, noticing is more likely to take place when there is a form-focused learning activity or specific corrective feedback involved[11]. In this regard, research on phonetic training with manipulation of perceptual cues (acoustic and visual) have shown that non-native listeners can improve their perception (Hazan & Simpson, 2000; Hazan, Sennema, Iba, & Faulkner, 2005). However, and assuming that noticing takes place either by means of external intervention or the learner's own devices, then the listener will be able to perceive that the sound in question, although still considered an L1 sound (as in options (b) and (c)), is somewhat more deviant from the usual L1 target values. Such noticing, however, will be rather elusive, as the warping effects exerted by the L1 categories will prevail over any attempt to change

[11]Unless the L2 learner has enrolled on a University course leading to the degree of translator or language teacher, it is very unlikely that the learner will receive any type of negative evidence and/or feedback regarding perception or pronunciation; any English language syllabus can confirm this assumption. The probability of receiving negative evidence in a more naturalistic learning environment is even smaller, unless the learner makes a mistake leading to the confusion of minimal pairs (and even so, the confusion must lead to a situation embarrassing enough to elicit any type of corrective feedback from the interlocutor!)

the number of categories and the consequent acoustic-to-linguistic mappings. Furthermore, it is expected that the previous categorisation attempts will feedback onto the perception grammar and leading to stabilisation of the system.

Finally, if the grammar takes option (d), then new acoustic-to-linguistic mappings will be set, so that a specific range of acoustic values becomes available to map to the new perceptual representations.

It is important to stress that the remapping and subset scenarios are indeed triggered by noticing, but that the potential inability to demote the relevant *Categ constraints would make L2 listeners to remain trapped in an 'IL loop': An in-between state where noticing has taken place and the perception grammar has become able to identify a less prototypical recurring value in the input, but has not been able to assign a specific target value and/or isolate it as a category in its own right, hence associating it to a native category in the form of a less prototypical member that eventually becomes either a member of a neighbouring category, or a subset of the L1 category onto which it is being mapped. As mentioned above, noticing seems to be a necessary, but not a sufficient condition to trigger category creation.

It becomes evident at this point that these options described will result not only in the different perceptual learning outcomes described above, but will also have different implications for the UG issue: is parameter setting a feasible option for perception? Can L2 speakers ever progress from L1-like to more target-like perception, and to what extent? If no corrective feedback is assumed, can L2 speakers access UG? Ultimately, such concerns will also have implications for a theory of learnability in L2 speech perception.

Moreover, optimal L2 speech perception poses a significant challenge for any attempt on modelling from an OT perspective. Since the Perceive and *Categ constraints are highly ranked, and since speech perception is understood as a categorisation process, then the only possible option for a token is to be assigned to the *existing* category that best fits the constraints of the grammar. From then on, the only possible way to bootstrap a different category is by means of distributional learning, so that a bimodal distribution (given by the tokens of /ʌ/ and /ɑ/) eventually breaks the boundaries of the L1 category. However, it seems that the set of tools given by the OT framework might not be enough to model such a process. Alternatively, the *Warp constraints could be promoted in favour of a perceptual outcome that allows for subcategorial learning; this requires the creation of in-between categories such as those proposed in Chapter 3 (Section 3.3.2).

2.7 Summary

Chapter 2 has presented a series of constraints that late-learners of English as an L2 face when acquiring the phonology of the language. Firstly, linguistic constraints (as we have been defined them) relate to the interactions between the specific elements of the native language and the second language. We have identified a phenomenon of transfer as one of the most important issues that hinders acquisition, particularly in the case of L2 English learners with Spanish as L1 and the difficulty in acquisition of the /ɑ/ - /ʌ/ contrast; this is due to the smaller vowel inventory of Spanish and the loss of contrast between these two vowels entailed by transfer of the L1 vowel system.

A second linguistic constraint that was identified is the input problem: Generativist approaches to SLA have stated the importance of the type of input that L2 learners receive. In this regard, the fact that obtaining corrective feedback in perception is unlikely, decreases the likelihood of noticing; this implies that learners remain unaware of the differences between the incoming input and the output that they produce.

Maturational constraints have also been identified by the field of SLA as relevant in the acquisition of a phonology, more than any other language domain. Importantly, the age factor seems to be crucial for the acquisition of a phonology, which is why theorisations such as the CPH have not been entirely disproven. Likewise, these theorisations have been formalised by generativist approaches, which explain the difficulty in the acquisition of an L2 phonology via the hypotheses of no access, or partial access to UG.

Additionally, a further layer of difficulty is given by the specific perceptual mechanisms underlying perception of speech sounds. Here we have mentioned categorical perception and the Perceptual Magnet effect as psychoacoustic sources of hindrance in the perception of L2 sounds: While the former shows that discrimination of speech sounds is poor when the sounds correspond to the same category, the latter suggests that prototypes of L1 categories act as a perceptual magnet for L2 sounds, thus making them difficult to discriminate.

The previous issues apply to this case study, which is the perception of the /ɑ/ - /ʌ/ contrast by L1 Spanish speakers of English. The role of the perception grammar in the input-to-intake is crucial, and target-like perception of the contrast requires changes in this perception grammar, which in its default setting becomes an impediment for noticing. This is given by mechanisms that are inherent to the speech perception process, and as such it seems that overriding the L1 settings is not an easy task for L2 learners. However, and despite the possible difficulties, I have outlined four possible options: full intake (creation of new phonemic perceptual categories), no intake (preservation of the L1 categories) and the two different options that result from the ‘IL loop’ (subsetting and deflection), which can be described as in-between states. Such learning outcomes will have their counterpart in phonological theory, which will be the topic of Chapter 3.

CHAPTER

THREE

L2 PERCEPTUAL CATEGORIES AND THEIR PLACE IN PHONOLOGICAL KNOWLEDGE

A veces lamento hablar en español: escuchado desde la otra orilla debe ser algo incomparable, lleno de chasquidos y latigazos, terrible carga de caballería de abiertas vocales, por entre un campo erizado de consonantes clavadas como estacas.

La jitanjáfora
Alfonso Reyes

3.1 Introduction

As mentioned in Chapter 1, the gap between SLA studies on L2 phonology and proper phonological theory is a large one, insofar as the former is mostly concerned with *performance* in L2, with research questions that are focused on aspects such as the production or perception of “difficult” sounds in L2, and results that seldom theorise about the *representations* in L2. On the other hand, the latter has paid little attention to L2 knowledge, despite the myriad of possible interactions between the two phonological systems at intra-speaker level that are worthy of exploration.

Regarding the difficulty in acquiring new perceptual categories in L2, Chapter 2 has explored the main constraints in the acqusition of new sounds. In this regard, several models have attempted to describe and predict acquisition in L2 speakers, given the different possible initial states. But as we have already mentioned above, none of these models concern the phonological nature of these sounds; rather, they predict perceptual behaviour based on the differences between the L1 and L2 inventories. Nevertheless, they are relevant for the purposes of this research as they set the foundations for L2 speech perception research.

In this chapter I summarise and analyse the main research in L2 phonology and speech perception, from specific findings to experimental approaches and general models. I will also present four possible outcomes in terms of the nature of the representations in L2 phonology, which will be the core hypothesis of this research. Finally, I present experimental designs that takes all the previous discussions into account, which will in turn become the theoretical and methodological foundations of experiments 1 and 2 as presented in Chapters 4 and 5.

3.2 Current L2 phonology models

This section presents the three most influential models in L2 phonology: The Speech Learning Model (SLM) (Flege, 1995), the Perceptual Assimilation Model (PAM (Best, 1995), and the L2 Linguistic Perception Model (L2LP) (Escudero, 2005). We also offer an analysis of these three frameworks in terms of their contribution to phonological theory and how their main tenets relate to our case study, namely, perception of the /ɑ/ - /ʌ/ contrast by L1 Spanish learners of English.

3.2.1 The Speech Learning Model (SLM)

The SLM model (Flege, 1995) is a framework consisting of a series of postulates and hypotheses about the ultimate attainment of L2 learning. According to the author, “regardless of whether the SLM is ultimately supported or disconfirmed, it serves as a useful heuristic for planning research" (p.238). Hence, the model itself is ultimately committed to suggest paths for research in L2 phonology, rather than to offer a thorough explanation for L2 phonology phenomena. However, the SLM also discusses -though in a somewhat loose way- the nature of the sound categories that L2 speakers create.

The SLM is formulated as a set of four postulates, which are supported by seven interrelated hypotheses. Both postulates and hypotheses are presented below.

3.2.1.1 Postulates

Postulate 1 states that “the mechanisms and processes used in learning the L1 system, including category formation, remain intact over the lifespan, and can be applied to L2 learning" (p.239). This postulate is a reaction to the Critical Period Hypothesis (CPH) which claims the existence of a maturational constraint in SLA, with the corresponding implications for ultimate attainment. The SLM rejects this hypothesis on the basis of an experiment showing a positive correlation between Age of Arrival (AOL) and perceived foreign accent in native speakers of Italian who had lived in Canada for an average of 30 years (Flege, Munro, & MacKay, 1995). Even though this can be seen as evidence in favour of the CPH, Flege’s argument is that the correlation is perfectly linear between ages 0-25, instead of showing a stationary trend after a certain age. However, this argument does not necessarily apply to the creation of new perceptual categories as a) it refers to speech production (and more particularly, perception of a foreign accent); and b) is made on the basis of a subjective judgement (made by NS nevertheless).

Postulate 2 states that "language-specific aspects of speech sounds are specified in long-term memory representations called *phonetic categories*" (p.239). Postulate 3 is closely related to 2, as it states that "Phonetic categories established in childhood for L1 sounds evolve over the life span to reflect the properties of all L1 and L2 phones identified as a realisation of each category" (p.239). These two statements lie at the core of the present research, insofar as it is the most explicit claim to date about the nature of L2 sound categories, and to which this research agrees in nature: A long-term memory representation which does not hold the same status as a phonemic category. Additionally, Postulate 3 suggests that only L1 categories are able to encode the phonetic information of newly acquired sounds.

However, it is worth noting that the concept of phonetic category has been created, as van Leussen and Escudero (2015) point out, from the perspective of isolated sounds, rather than nonnative contrasts (p. 2). The concept of phonetic category, although appealing for the purposes of this work inasmuch as it makes a specific claim (albeit somewhat indirectly) about the nature of the categories that L2 speakers create, seems to better apply to consonants than to vowels. However, Flege still transposes the concept to the case of L2 vowel perception, even though it seems to have been formulated in order to explain the results of experiments regarding VOT stop values in bilinguals (Mack, 1990; cited in Flege, 1995). In fact, the research cited by Flege is not conclusive with regard to whether L2 speakers are able to create vowel categories at all. Firstly, the experiments used concerned either oddball paradigms or dissimilarity ratings which elicited behavioural responses. Secondly, perceived dissimilarities between two TL vowels were found to be significant only when vowels were further away in the acoustic space, but not when they were adjacent. And lastly, the experiments did not considering the perceptual mappings that take place during the first stages of L2 learning, where all sounds are perceptually assimilated to an L1 vowel, as Flege suggests.

There are several reasons why a category creation model based on consonants might not be suitable for vowel sounds. Firstly, and from a production viewpoint, the articulatory gestures involved in consonant production produce much less dispersed acoustic values than vowels: while the former are produced by articulatory gestures that require more constriction in the vocal tract, vowels are characterised by the free passage of air which leads to less strict specifications and more dispersion of their acoustic values (Fry et al., 1962). Secondly, from a perceptual viewpoint, Pisoni (1973) has shown that the perceptual mechanisms used for vowel discrimination are not the same as those used for consonant sounds; for instance, it is stated that short-term memory plays a fundamental role in perception of vowel sounds. And thirdly, it has been pointed out that perception of vowels is gradient, rather than categorical (Schouten et al., 2003).

Finally, Postulate 4 posits that “Bilinguals strive to maintain contrast between L1 and L2 phonetic categories, which exist in a common phonological space” (p. 239). This postulate has two implications. The first is related to the fact that it is assumed that bilinguals would prefer to maintain a contrast, even though another plausible choice is that L2 speakers choose to decrease the amount of functional load by simply keeping their L1 categories and assuming that two minimal pairs in the L2 are homophones. However, it seems that speakers do not necessarily decide between one and another, but could also experience some kind of ‘in-between’ learning, of which little is known or has been theorised about. The second implication is that if these categories exist in a common space, then it is not impossible that even though some learning can take place in L2 speakers, these new categories would be somehow still related to L1 categories.

3.2.1.2 Hypotheses

The SLM postulates presented above derive from seven hypotheses, all of them supported by findings of previous research.

Hypothesis 1 (“Sounds in the L1 and the L2 are related perceptually to one another at a position-sensitive allophonic level, rather than at a more abstract phonemic level", p. 239) derives from the first two postulates. This hypothesis has an important implication: L2 categories do not have a phonemic status, but they are somehow incorporated into the L2 speaker’s grammar. This statement is supported by research on the contrast between English consonants /ɹ/ and /l/, which is difficult to perceive and produce by L1 speakers of Japanese (Strange, 1992; cited in Flege, 1995). However, perception of these two consonants changes according to the context: Speakers are considerably more sensitive to the phonetic differences when they are in word-final position. But again, a possible reason for this asymmetry is that speakers of Japanese also categorise word-final /l/ as either “ru” or “o”, whereas word-final /ɹ/ was categorised as “a” (Takagi, 1993; cited in Flege, 1995). This in turn would not mean that L2 speakers are creating phonetic categories that are somehow independent of the L1; rather, perception of the contrast is explained by perceptual assimilation to *different* L1 categories.

Hypothesis 2 states that “A new phonetic category *can be established* for an L2 sound that differs phonetically from the closest L1 sound if bilinguals *discern at least some of the phonetic differences* between the L1 and L2 sounds” (p. 239)[1]. This principle seems to be accurate, particularly in the light of later studies such as that of

[1]The emphasis is mine.

Escudero and Boersma (2004). In this study, perception of the /i - ɪ/ contrast can be learned by L1 speakers of Spanish by either perceiving F1/F2 differences or duration. More interestingly, such weighting of the available perceptual cues is learned from the environment: NNSs will learn how to weight the cues in a similar proportion to what NSs of the specific TL variety do. However, if the only perceptual cue that is available for the contrast (e.g. formant values) is also the same one that NNSs with fewer L1 categories have, then the ability to discern these phonetic differences between L1 and L2 categories are lower. Crucially, Hypothesis 2 emphasises the importance of noticing such differences in input, which, as discussed in Chapter 2 may not necessarily lead to intake.

According to Hypothesis 3, "the greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned". Evidence of this is Takagi's research mentioned above, which, according to Flege, leads one to believe that it is more likely that Japanese speakers of English will discern the phonetic differences between the English /ɹ/ and the Japanese /r/, than between the latter and the English /l/. However, this argument fails to see that, just as Hypothesis 1 explains, the reason why these phonetic categories can be formed in this case is not because L2 speakers are able to discern the phonetic differences of the sounds *per se*, but because these phonetic differences are linked to the perception of two different representations in the L1.

Hypothesis 4 states that "The likelihood of phonetic differences between L1 and L2 sounds, and between L2 sounds that are noncontrastive in the L1, being discerned decreases as AOL increases" (p.239). This hypothesis is based on the study by Flege et al. (1995) mentioned in Postulate 1, which is supported by subjective judgements given by NSs and no actual acoustic comparison between the average acoustic values of TL forms and those found in the IL. However, Flege does note that the lack of ability to discern between two different categories as AOL increases may arise at different levels of processing, and not necessarily as a consequence of sensorineural loss: Given that automaticity in speech perception is gained during childhood (Linell, 1982; Mayberry and Eichen, 1991; cited in Flege, 1995), it is likely that adults disregard certain phonetic cues in the L2 input in order to assign more resources to higher order processes in the L2 (p. 241). However, and should this be the case, it is also possible that the regression results obtained may show a broken line (if performing a different type of regression, such as LOESS), or a flatter slope in the case of a regular linear regression. All in all, the results might not be explained specifically by AOL itself but by the amount of experience in the L2, and it would also be safe to assume that according to Flege's view, late L2 speakers do not create *phonemic categories* that are similar to those of the L1, if they create any category at all. This is as far as defining an L2 category goes, not only in Flege's framework but in the L2 phonology literature in general.

According to Hypothesis 5, "Category formation for an L2 sound may be blocked by the mechanism of equivalence classification. When this happens, a single phonetic category will be used to process perceptually linked L1 and L2 sounds (diaphones). Eventually, the diaphones will resemble one another in production" (p.239). Again, this statement explains the way in which L2 speakers with different VOT values in their L1 manage to resemble the prototypical values in the L2 (Flege, 1987). The study considered NS of English who were L2 speakers of French and vice versa, and measured VOT values for /t/ in utterances in both languages. The results showed that more experienced L2 speakers were closer to the prototypical values in the TL, but not completely: their VOT values were still shorter than those of native speakers of English. Furthermore, these experienced L2 speakers also showed an effect of L2 experience in the L1, as their /t/-initial utterances in French showed longer VOT values than those of monolingual speakers of French. Similar results were reported by Flege and Eefting (1987) regarding VOT values in native speakers of Spanish with high proficiency in English. However, VOT is also a predominantly time-related feature, just like vowel duration; formant values, on the other hand, are predominantly frequency-related. It is expected then that L2 learners may show different behaviour when dealing with duration than with formant frequencies, given the different nature of the cues.

Hypothesis 6 states that "The phonetic category established for L2 sounds by a bilingual may differ from that of a monolingual if: 1) the bilingual's category is 'deflected' away from an L1 category to maintain phonetic contrast between categories in a common L1-L2 phonological space; or 2) the bilingual's representation is based on different features, or feature weights, than a monolingual's" (p. 239). It can be assumed that this deflection is given by a recategorization of two L2 sounds that in their initial state are perceived as instances of one L1 sound, in which case the creation of a phonetic category is being partly aided by the L1 and partly hindered by it as well, as it can also be assumed that deflection of the L2 token onto *another* L1 category would also result in non target-like perception or production insofar as the L2 category does not have the same prototypical values as any L1 category. With regard to 2), studies like Flege (1987) or Escudero and Boersma (2004) reinforce this statement.

Finally, Hypothesis 7 posits that "The production of a sound eventually corresponds to the properties represented in its phonetic category representation". This statement, as weak as it is, can be interpreted as a way to commit to a Full Access theory of a final state in the acquisition of L2 phonology, which, as far as the evidence shows, remains unclear.

In sum, Flege's approach is not necessarily conclusive regarding the creation of phonetic categories. In this respect, it is worth noting that the concept of phonetic category used by Flege is not particularly clearly defined as it does not address the problem of the phonological nature of categories by what it is; rather, it is defined by what it *could* be. Specifically, the SLM does not make any specific claims in cases of 2-to-1 L2-to-L1 mappings, which certainly poses a problem for late-learners of an L2. Although the model predicts that late-learners will present more difficulties in perceiving the contrast, suggesting that they are unable to learn new categories would contradict the first postulate of the SLM, according to which the language learning mechanisms remain intact over the lifespan. This research would wish to carry out further investigations into this problem and understand the extent and nature of this type of learning.

3.2.2 The Perceptual Assimilation Model (PAM)

The Perceptual Assimilation Model (PAM) (Best, 1995), is a direct realist approach to naïve cross-language speech perception. It discusses the main mechanisms of human speech perception in general, and also presents a set of predictions (based on empirical observations) regarding the perception of both isolated sounds and contrasts in nonnative speech.

PAM drifts away from the classic psychoacoustic approach to speech perception, which assumes that the perceptual primitives are proximal acoustic cues that are used to retrieve information stored in mental representations. Rather, PAM assumes an Articulatory Phonology view, according to which the basic raw material to be perceived are not acoustic cues but distal articulatory gestures. An important implication of such a speech perception model is that it assumes that the perception of speech sounds is not specific to human speech. In fact, PAM posits that acoustic and articulatory features of human speech (i.e. what is *actually* being perceived) are not a means to reach mental representations: They are, in fact, the representations themselves.

Additionally, the main prediction made by the model (and that does not necessarily depend on the direct realist approach to work) is that cross-language speech perception can be understood as a *perceptual assimilation* phenomenon where the nonnative segments are mapped onto L1 representations according to their similarities to the gestural constellations of the native language (p.193). This is a feature that the SLM does not have: The ability to make predictions with regard to nonnative contrasts, and the link to such perceptual outcomes with the specifics of the native language system.

PAM mentions six possible assimilation types for nonnative contrasts (p. 195):

- Two-Category Assimilation (TC type), where two nonnative categories are perceptually assimilated to two different native categories. Discrimination is expected to be excellent.

- Category-Goodness Assimilation (CG type), where two nonnative segments are assimilated to the same native category but with different levels of category goodness. Discrimination is expected to be moderate.
- Single-Category Assimilation (SC Type) where both nonnative sounds assimilate to the same native category but are equal in terms of category goodness. Discrimination here is expected to be poor.
- Both Uncategorizable (UU type) where nonnative sounds are perceived as speech sounds, but the listener is unable to assimilate them to any native category. Discrimination is expected to range between poor to very good, depending on perceptual distance and/or potential proximity to a native category.
- Uncategorized versus Categorized (UC type): One sound is assimilated to a native category and the other is not. Discrimination is expected to be good.
- Nonassimilable (NA type): the two sounds are considered to be nonspeech sounds. Discrimination is expected to be good.

One of the problems of an Articulatory Phonology model is that it bases the study of perception on constellations of gestures as input for the listener, instead of the usual acoustic cues. Furthermore, the fact that it denies the presence of mental representations makes a theory of competence in L2 nonexistent.

A second problem of the model is the shortness of its scope with regard to the learning process. The model does not make any claims regarding a final state in L2 speech perception, and/or the learning of an L2 sound system in general; naturally because it is a model of naïve non-native speech perception, and not one of perceptual learning. Despite having a great predictive power, the lack of explanation regarding how perception changes throughout the process of learning an L2 makes the assessment of the model from the perspective of a theory of language learning impossible.

3.2.3 The L2 Linguistic Perception Model (L2LP)

The L2LP model (Escudero, 2005) is a major contribution in the field as it links phonological theory and speech perception, which has been left outside the linguistic domain and therefore considered to be only a phonetic or psychoacoustic phenomenon (p. 9). The model is based on Functional Phonology (Boersma, 1998), which models speech perception and production within an OT framework with the addition of a stochastic component in order to account for different phonological processes. Escudero's main hypothesis is that an L2 speaker -provided with rich input- will move from a Full Copying state to a Full Access one (Escudero & Boersma, 2004, p. 553)

One of the most important assumptions of the L2LP model is that speech perception includes two main components: a *mapping component* (a perceptual grammar) and a *representational component* (the perceptual categories) (p. 42). This twofold model is crucial for the purposes of this research, as the focus of this work is on the likelihood that perceptual categories are created from a perceptual input that corresponds to two different categories in the L2, but only one in the L1. The importance of such a distinction is that it defines perception as a language-specific phenomenon, as different languages create their perceptual categories on the basis of different acoustic values. For instance, even though both Canadian English (CE) and Canadian French (CF) have the /ɛ/ and /æ/ categories, the boundaries between the categories differ in terms of their specific values. While CE shows a perceptual boundary between /ɛ/ and /æ/ categories at 756 Hz, the boundary in CF is at 637 Hz; hence, any token whose acoustic values lie within these two boundaries will be categorised differently by native speakers of CF and CE. (Escudero, 2005, p. 255).

The type of constraint that the L2LP model proposes for adult speech perception is an auditory-to-phonological one, with raw acoustic values (e.g. F1 = 300 Hz) at the acoustic end and a phonological category (e.g. [+high], or /i/ for the purposes of the /i-ɪ/ contrast in English) at the phonological end. Hence, a constraint in this model would be interpreted as "300 Hz is not /ɪ/".

The L2LP model considers five 'ingredients' for optimal L2 speech perception. The first ingredient is the description of optimal L1 and L2 speech perception. Escudero defines optimal L1 speech perception as the best possible way to perceive sound categories in the learner's first language, as they have optimal perception grammars. This is refered to as the *optimal perception hypothesis*, according to which "listeners maximise their probabilities of understanding speakers by making decisions that match their intended message" (Escudero, 2005, p. 88). Target L2 optimal perception is "the best *possible*[2] way of perceiving the learner's target L2, which is predicted to be found in native speakers of the target language" (p. 89). The description of an optimal L1 and L2 grammar takes into account the specific values in the acoustic signal that lead to the perception of a certain category, and the boundaries between categories. These values will then lead to the creation of constraints with a specific range of values for a category.

The second ingredient refers to the initial state and its corresponding description (p. 97). The L2LP model assumes that an initial state is that of optimal L1 perception, with the reference acoustic values and constraints from the L1 being fully transferred onto the L2. The linguistic explanation for this state is given by the Full Transfer hypothesis (Schwartz & Sprouse, 1996), which proposes that all the representations of the L1 are copied at the onset of L2 acquisition. The L2LP model assumes a Full Copying hypothesis for the initial state, according to which the

[2]The emphasis is mine.

learners will reuse their perceptual mappings of the L1 (both the representations and the grammar) in order to perceive the sounds of a second language. The rationale for proposing such an initial state is rather simple: it is better to perceive something than not to perceive anything at all (p. 100-101). In this sense, copying the L1 perceptual grammar and its system of representations ensures that every sound in the L2 is assigned to a perceptual category insted of leaving it unparsed.

A third ingredient is the L2 learning task (p. 105). In this respect, and as mentioned above, Escudero claims that there are two different learning tasks: a **perceptual** task and a **representational** task. In the case of 2-to-1 mappings (or Single Category assimilation, according to Best's framework), where the stimuli received in the L2 corresponds to two different categories with their acoustic values being mapped onto the one same L1 category, the learning task involves both a perceptual task and a representational task. The representational task consists of the splitting of L1 perceptual mappings: The learner will create two perceptual targets out of one single representation. The perceptual task, consequently, consists of learning the range of values that are ought to be mapped onto the corresponding created representations.

The fourth ingredient is L2 development (p. 109). This refers to the mechanism that takes place in order to move from an initial to a final state. The model hypothesises that the mechanism that ensures development of optimal L2 perception is the same as that found in L1 perceptual development. This mechanism is the Gradual Learning Algorithm (GLA), which was first introduced by Boersma (1997) and proposes a mechanism that assumes category creation in L1 essentially as distributional learning, where categories are formed from the perceptual input given by the environment. However, this learning is also regulated by other aspects, such as the frequency in which a category shows given a certain context. Crucially, the GLA models the learning process by proposing an algorithm that causes promotion and demotion of constraints: for instance, a listener can learn how to deal with categorisation by continuous reassessment of the constraint order according to the outcome. If the categorisation was not correct, then the constraints that generated such perceptual outcome will be rearranged. Promotion and demotion of perceptual constraints would then allow an L1 learner to move from an unbiased state –where the main criterion for categorisation is the acoustic closeness of the input value to the center values of a given category– to a new one where constraints reorganise in favour of the category intended by the speaker; subsequent iterations of this learning process with constraint demotion and promotion along a continuous scale ensure optimal perception.

Likewise, GLA can model perception of non-native categories by reranking the relevant constraints: according to L2LP, an L2 listener should be also able to rearrange her constraints in order to achieve optimal perception in L2. The L2LP model also mentions the role of feedback obtained from the lexical domain, which requires that the L2 learner *notices* the mismatch between their underlying representation and the actual input; hence, if the learner's UR for both 'ship' and 'sheep' is /ʃip/, and the input has the form of 'there is a [ʃɪp] in the field' then the learner will either: (a) correct a potential mapping to 'ship' onto 'sheep' and correct it after the rest of the sentence is given; or (b) have both possible meanings in mind and choose 'ship' only after the rest of the sentence has been given. I have mentioned my reluctance in Chapter 2 to accept such a hypothesis for category splits, since noticing may not necessarily take place.

Finally, the fifth ingredient of the model is the L2 end state (p. 113). Predictions for the model in regard to perception rely largely on the input amount: The richer the input, the more likely it is that the learner achieves optimal, target-like L2 perception, with L1 perception remaining optimal as well. This prediction is linguistically explained by two hypotheses: (a) that rich input is enough to override the effects of a potential small cognitive plasticity; and (b) that the perceptual grammars are separate systems, thus ensuring its stability. Escudero describes plasticity from a computational point of view, equating it to the learning steps that the GLA describes. While an L1 learner takes bigger learning steps with more mistakes along the learning process, an L2 learner takes smaller learning steps, but also makes fewer mistakes. By rich input, the model means that it must be as abundant as the one received when learning the L1, but that it must also contain negative evidence. The hypothesis of having two separate perceptual grammars is supported by Grosjean (2001) and his notion of *language mode*, according to which an L2 speaker activates different perception and production skills according to the parameters of the language that is being perceived from the environment. For instance, in order to obtain language-specific responses in categorisation of stops /b-p/ and /t-d/, Antoniou et al. (2012) triggered different language modes in Greek/English bilingual speakers by conducting the experiments entirely in either Greek or English (p.585), which resulted in the same speaker providing different categorisations for the same segment.

The L2LP model takes acquisition of native-like L2 categories mostly as a function of input amount. However, it does not necessarily follow from this that the input itself is sufficient in order to achieve Full Access. Furthermore, the assumption that the optimal L2 speaker is able to create native-like categories on the basis of input amount regardless of their age or any other cognitive factors challenges the fact that the L1 input received by the speaker before learning the L2 is too robust to be overridden by the new language. In this regard, it is worth remembering that according to the ASP model mentioned in Chapter 2, L1 phonological perception is a robust, over-learned process that speakers cannot easily reverse. In addition, the model's main evidence for the creation of new perceptual categories in L2 is an experiment made on the basis of labelling only, though it does give excellent evidence regarding the use of perceptual cues by L2 listeners.

Additionally, the model posits that the L1 and the L2 are separate perceptual grammars. According to the L2LP model the initial state in L2 learning is a blank slate to which the perceptual grammar of the L1 is copied. However, in a case such learning the /i - ɪ/ contrast by native speakers of Spanish, the learner still has a new perceptual cue to rely on when creating the new categories, which is not always the case when learning contrasts: Duration does not necessarily have to co-occur with different F1/F2 values.

Finally, the L2LP model explains the learning of a perceptual grammar, but it is not entirely clear how the grammar operates with regard to the creation of new categories: Why does L2LP allow for the generation of new perceptual categories as candidates? Rather, what is the reason for GEN to propose a candidate such as /ʌ/, if the L1 grammar can still handle the perceptual input by making an existing candidate win? I have mentioned in Chapter 2 that the addition of constraints from earlier formulations of perceptual OT may offer a solution; this will be discussed further in Section 3.3.2.

3.2.4 Interim summary

Table 3.1 summarises the main divergences among models. The most important remark to be made is that all of these models focus on different aspects of L2 speech perception. The SLM model deals with ultimate attainment in L2 phonology, with an emphasis on the type or representation that L2 speakers are able to create, and the corresponding variables that influence ultimate attainment (where AOL seems to be the crucial one). The PAM model has a completely different view of what speech perception is, as it advocates for a direct realist model where representations (if there are any) are learnt from observations of auditory gestures rather than raw acoustic values. However, the model constitutes a major contribution as it is the first model that considers perception of nonnative contrasts as a main feature of L2 speech perception, as well as it predicts perception on the basis of the relation between L1 and L2 categories. This point of view naturally implies that the initial state of L2 speech perception is the category inventory of the L1, thus advocating for a Full Transfer hypothesis in the phonology domain. And finally, the L2LP model is a thoroughly formulated, highly explanatory model of the trajectory described by L2 learners between an initial and a final state, although it does not deal with the status of the categories that learning of an L2 perceptual grammar may be forcing to create; or how this grammar operationalises the qualitative changes undergone by these new categories.

However, the models coincide in that the initial state must be that of the L1 in terms of the categories that the input is being mapped onto. Even though we adhere

Table 3.1: Models of L2 speech perception

Model	Learning stage focus	Nature of perceptual category	Initial state	End state
SLM (Flege, 1995)	Ultimate attainment	Phonetic category	Blockage due to equivalence classification	Potential creation of phonetic categories
PAM (Best, 1995)	Naïve nonnative perception	Articulatory gestures	Perceptual assimilation of nonnative sounds onto similar native sounds	Not specified
L2LP (Escudero, 2005)	Ultimate attainment	Perceptual category	L2 perceptual categories merged into one native perceptual category	New perceptual categories created by learning of TL-like perceptual cue weighting

to many of these models' main points regarding the initial state, they do not seem to provide answers to the problem presented in Chapter 1, namely, the final outcome of category redistribution (if it happens), and the mechanisms that implement such changes. Furthermore, these models do not include a discussion from a cognitive point of view with regard to the substance of the representations, or the way they conceptualise the categorisation process in general. In the following section a brief discussion on category creation in L2 is offered, along with the potential learning outcomes.

3.3 Category creation in L2

Creating perceptual categories in L2 is not similar to any process that could happen in the L1, neither during L1 acquisition nor at a later stage. Whereas the L1 phonology acquisition process starts form a complete blank slate, the L2 phonology acquisition process is characterised by a non-zero initial state: Acquisition of L2 phonology begins from the basis of a previously acquired phonological system. Additionally, this process is different to anything similar in the L1 given that phonological systems are not expected to change during the person's lifespan in terms of adding a complete new phonemic category[3]. Hence, and although other domains

[3]Although cases of splits in phonology due to processes of language change exist, they are expected to originate from allophones of a certain vowel phoneme, thus appearing first only in different phonological contexts. A well-known example is the [ð] - [θ] allophonic pair (both allo-

of language knowledge such as the lexicon, the morphology, and syntax do change during the lifespan, it is not expected that a specific speaker would have to learn a new perceptual category for a new sound category in their L1.

Additionally, learning an L2 implies, as seen in Section 2.2.1, a process of full transfer during the initial state. This is an important feature of L2 phonology: Perception is, at least initially, aided by the categories of the L1, or in Trubetzkoy's words, "sieved" through the grid of the L1. Hence, to the eyes of the L2 learner, L2 perception is L1 perception of non-prototypical sounds. Eventually, and if we agree with SLM and L2LP, perception changes towards an L2-like perception, both in terms of the categories that are available for the incoming input and the mechanisms that deal with the raw acoustic signal.

Creation of new perceptual categories can present different levels of difficulty, according to the relation between the L1 and the L2 phonological systems. Regarding perception, Best (1995) presents six different types of assimilation of contrastive sounds in the TL (as seen in Section 3.2.2), and Escudero (2005) presents three different initial state scenarios: learning *new* L2 sounds (via splitting an L1 category, as in Spanish L1 learners of SBE); learning *subset* L2 sounds (via merging two L1 categories, as in Dutch learners of Spanish); or learning *similar* L2 sounds (via rearranging the perceptual grammar to different reference ranges and hence readjusting the boundaries, as in English L1 learners of CF).

The focus of this work is on what Best defines as Single Category Assimilation (SC Type) and what Escudero calls a *new* scenario. This type of perceptual phenomenon is linked to an L2 that has *more phonological categories than the L1*, which will necessarily imply an initial mapping of two or more L2 categories onto one L1 category. The learning process should move towards a perceptual split, which has two main implications: firstly, that the learner will have to notice that what they perceive as one sound is actually a misperception of two sounds; and secondly, that learners will have to bootstrap themselves out of that too-broad perceptual category by learning new perceptual strategies and hence creating new categories.

However, learning a new sound category (production- or perception-wise) is a process that is subject to certain effects that base their origin in the early stages of L1 learning. According to the PME/NLM, categories are created as early as 6 months old (Kuhl, 1991; Kuhl & Iverson, 1995; Kuhl et al., 1992). They are also deemed to be language-specific, as infants who are around 10 months old are already unable to perceive differences between a native-like stimulus and a non-native-like stimulus that is acoustically close to the native one. As the PME/NLM model posits, these categories have a very strong effect on perception of linguistic sounds and offer a very robust warping effect. Therefore, if the new L2 input offers tokens of a new phonemic category that is similar to an L1 category, the expected effect is

phones of /θ/) in Old English, which became contrastive units in Modern English after the loss of the conditioning environment (Hoenigswald, 1960; Ringe & Eska, 2013).

that the learner will perceive these tokens as members of the L1 category.

Despite the fact that the L1 might block the contrasts of two L2 categories due to the mapping mentioned above, the acquisition of such contrasts might also become an optional process, provided that such a lack of contrast would not affect day-to-day communication. This scenario can be possible if the L2 does not provide any minimal pairs; and if they do, they either correspond to different word categories (e.g. a noun and a verb) or are lexical items that would rarely appear in the same context: if we can imagine a language with the words 'camel' and 'spoon' being minimal pairs that an L2 learner would perceive as the same, then it is still unlikely that the learner would need to learn the contrast for everyday communication.

The reasons mentioned above show the importance of this study in terms of its implications for learnability and the development of a phonological account of L2 acquisition. It is not enough to know which categories pose more difficulties for learning; it is also important to understand the learning process that takes place (if any), and the outcome of this learning in terms of the specific characteristics of the perceptual categories that are being learnt. Is this learning robust? How reliable is it when compared to L1 perceptual categories? If we consider that perception is mostly a categorisation process, then it is worth looking into the creation and nature of such perceptual categories. If there are long-term representations, are they still overlapping with, or remaining a subset of, an L1 category? Does the nature of these representations affect categorisation? What is the role of discrimination in predicting a categorisation outcome in L2? Are these categories weaker than those of the L1 due to the strength of L1 categories, which at the same time create a strong warping effect? These are only some of the questions that can be posed, to which I expect to offer some answers.

3.3.1 Previous studies in L2 vowel perception

Previous work in the perception of L2 vowel categories has shown somewhat contradictory results, although they all agree with respect to the role of the L1 inventory. However, it seems that it is not the inventory *per se* that generates difficulty in L2 speakers to perceive non-native contrasts: Language-specific cue-weighting, feature availability, and single-category assimilation are counted among the possible factors affecting vowel perception.

3.3.1.1 Acoustic-based approaches

Fox, Flege, and Munro (1994) performed a multidimensional scaling (MDS) experiment where Spanish- and English- speaking subjects were asked to rate the dissimilarity of all the possible vowel pairs containing Spanish vowels /i/, /e/, and /a/, and English vowels /i/, /ɪ/, /ɪ/, /eɪ/, /ɛ/, /æ/, /ʌ/, and /ɑ/. The results showed that Spanish-English bilinguals would present a clustering of vowels in a way similar to monolingual English speakers, and that vowel height seemed to be the most important dimension used by them in order to give their ratings. In a similar study with dissimilarity rankings, Flege, Munro, and Fox (1994) discovered that experience in the L2 did not play a significant role in perceiving dissimilarity: Both inexperienced and experienced groups of L1 Spanish speakers of English gave similar dissimilarity rankings to the same vowel pairs. At the same time, a universal component was found: Dissimilarity rankings increased as the vowel pairs were further apart in the acoustic F1/F2 space. The experiment considered an additional oddity task with triads of vowels that were either adjacent or non-adjacent in the perceptual space, showing that accuracy was higher in triads consisting of vowels that were given a high dissimilarity rating in the previous task.

Iverson and Evans (2007) investigated the effect of the L1 inventory on L2 speech perception by analysing three acoustic cues: perception of formant targets, formant movement, and duration. The study was performed on subjects with different language backgrounds: Spanish, French, German, and Norwegian. Perceptual tasks developed and carried out by the subjects were identification of English vowels in quiet and noise, L1 vowel identification in noise, L1 assimilation, and goodness ratings. Results showed that speakers with bigger L1 vowel inventories were more accurate in identification, but that the use of perceptual cues did not differ greatly among groups. However, even though different tasks were performed more accurately by L1 speakers who had either bigger inventories or gave prevalence to similar perceptual cues, all groups suffered from difficulties in the identification task in absence of a perceptual cue (i.e. manipulation via formant flattening, equalised duration) to the same extent.

Escudero et al. (2009), on the other hand, showed that cue-weighting is language-specific, and hence it does affect the perception of L2 contrasts, with the consequent implications for perceptual category learning. This research considered listeners of L1 Dutch, L1 German, and L1 Spanish. Participants were asked to carry out an XAB task (ISI = 1.2s), with stimuli corresponding to a F1/F2 - duration continuum between the Dutch vowels /aː/ - /ɑ/[4]. While Spanish only has the L1 vowels /a/ and /o/ as the closest native categories, German has a vowel contrast similar to Dutch but reversed in terms of duration: /a/ - /ɑː/. Hence, since /ɑː/ is long in German, the short Dutch /ɑ/ is more likely to be perceived as a German /ɔ/. The results showed that L1-Spanish speakers of Dutch are no less accurate in categorising this vowel pair; in fact, they even showed the same bias towards /ɑ/. German speakers were in general less accurate than L1-Spanish speakers, and did not show

[4]While Escudero et al. assume that there is an /aː/ - /ɑ/ contrast in German, this is controversial as it has been usually claimed that the contrast is just given by vowel length /aː/ - /a/, and not vowel quality (cf. Wiese, 2000). However, the authors of the study also mention that the spectral values are largely overlapped and that the main difference between these vowels is in fact duration (p. 454).

the bias mentioned above. Regarding cue-weighting, the experiment showed that L1-Spanish speakers rely more on duration, even though they do not have a vowel-duration contrast in the L1. The authors offer three possible explanations for this phenomenon: As the result of either a stage in L2 acquisition; as a result of universal saliency of duration; or as a result of L1 transfer, according to which allophonic vowel duration is being transferred from the L1.

Bohn and Flege (1997) tested production and perception of the new vowel /æ/ in two groups of L1 German speakers, one with advanced English language experience and another with little experience. Subjects were asked to categorise vowel sounds obtained from a /ɛ-æ/ 11-step continuum. The continuum was repeated with different duration values: 150 ms, 200 ms, and 250 ms. The results showed that the inexperienced group used duration as a cue, while the experienced group showed more use of the spectral cue. However, the distribution curve showed by experienced speakers was still different from that of native speakers of English. In addition, the study concluded that perception and production do not seem to be correlated, as production showed to be much more target-like than perception.

In sum, acoustic approaches highlight the importance of the L1 phonological system in perception, and more importantly, that experience in the L2 does not seem to be necessarily linked to a major improvement in the perception of certain non-native contrasts. One remarkable exception is that of Escudero and Boersma (2004) which proves that L1 Spanish speaker of English learn how to use perceptual cues in a way similar to that of L1 English speakers in perception of the /i - ɪ/ contrast, although not entirely: Speakers also showed cue/weighting strategies that were neither L1-like nor L2-like.

3.3.1.2 Feature-based approaches

Feature-based approaches to L2 perception are remarkably few, but more importantly, they present contradictory results. Brown (1998) offers a feature-based approach to speech perception in the L2, claiming that the lack of a specific feature in the L1 impedes acquisition of non-native contrasts in the L2. Brown tested perception of /l-r/ by native speakers of Mandarin Chinese and Japanese, and predicted that given that Mandarin has the [coronal] feature contrast in a different pair of consonants, then they should have less difficulty in acquiring the /l-r/ distinction. On the other hand, Japanese speakers should present more difficulty in acquiring this contrast, given that the [coronal] feature is completely absent in the grammar of Japanese. The results confirmed the predictions, which has important implications for L2 acquisition theory: the L1 is an important factor that hinders acquisition of a non-native phonology.

Similarly, Pajak and Levy (2014) tested perception of non-native *consonant* length contrasts in native speakers of languages with *vowel* length contrasts: the goal was to test whether familiarity with the duration cue in certain segments could be transfered to a different major class. Speakers of Mandarin (sibilant-experienced), Cantonese, Korean and Vietnamese (length-experienced) were asked to discriminate in a same-different AX task minimal nonce word pairs modelled after Polish phonology, with long and short sibilants. While length is a highly informative cue for both vowels and consonants in Korean, it is less informative in Vietnamese and Cantonese and it only affects vowels. Finally, Mandarin does not show any length contrast, but it does show a contrast regarding place of articulation in sibilants. The authors chose Polish phonology for their nonce words because of its length contrast in consonants[5] and its place contrast in sibilants (alveolo-palatal/retroflex). While this place contrast is not exactly the same as the one found in Mandarin sibilants, the similarity in the spectral cues that allows speakers to distinguish them made Polish a good candidate for testing the sibilant contrast in Mandarin speakers (p. 151). Sensitivity measures (d-prime scores) showed that Korean speakers performed best, followed closely by speakers of Vietnamese. Cantonese speakers were the group with the worst performance of the length-experienced languages (in Cantonese, length co-occurs with other perceptual cues, which explains the difference in performance). However, the three length-experienced groups were still ahead of the Mandarin group, which had the worst overall performance. These findings are in line with Brown's feature-based approach, insofar as they advocate for an L2 speech perception model that does not only take perceptual warpings into account, but also sensitivity to perceptual cues that correlate with a more abstract phonological feature. Furthermore, the authors claim that the level of informativity of a certain perceptual cue in the L1 causes a gradiency effect in L2 perception; that is, that perception of a non-native contrast in L2 will depend on the weight that the L1 assigns to the perceptual cues involved in the perception of such contrast.

However, further research on non-native perception of vowel sounds produced different results. Barrios et al. (2016) pose the question whether feature availability in the L1 is a necessary condition to achieve native-like perception of an L2 contrast in vowel segments. Barrios et al. investigated the role of phonological features in perception of two contrasts that are present in English but not in Spanish: /æ/ - /ɑ/ and /i/ - /ɪ/. The study hypothesises that since the [± back] feature is also present in the L1, then L1 Spanish speakers who learn English as their L2 should present better discrimination between /æ/ and /ɑ/ than between /i/ and /ɪ/, as the latter presents a tense/lax contrast that the Spanish vowel system does not have. The experiment considered two subject groups: one consisting of native speakers of English, and the second one composed by L1 Spanish speakers with a high proficiency in English. The tasks consisted of an AX discrimination procedure, with an ISI of 1,500 ms,

[5]The authors specify that the length contrast in consonants is a consequence of two different processes: borrowings from different languages, and morphological processes. Therefore, it is not a length contrast in the traditional sense of the concept.

and a medium-term repetition priming paradigm, with English words and nonce words. However, the results showed that even though experience in the L2 played a significant role in increasing accuracy, non-native speakers of English had the same degree of difficulty with the /æ/ - /ɑ/ contrast as with /i/ - /ɪ/. The study concludes that feature availability is neither a sufficient nor necessary condition in order to acquire a non-native vowel contrast.

Thus, feature-based approaches seem to show that feature availability in the L1 does not necessarily apply in an across-the-board manner: while it seems to be a significant factor in consonant discrimination, it does not appear to be an important factor in vowel perception.

3.3.2 Possible learning outcomes

The discussion in Chapter 2 showed how the acquisition of two L2 perceptual categories from an old L1 category poses a learning problem in terms of its feasibility and complexity; in this regard, one of the L2 phonology models in the previous review predict category learning at a phonetic level (SLM), while a second model predicts that L2 learners acquire native-like cue-weigthing patterns that allow them to create these new categories. However, one crucial question remains: What is the *nature* of the perceptual categories acquired by L1 Spanish late-learners of English, and what is the most likely learning outcome in speakers that have reached a high level of proficiency? I claim that four learning scenarios are possible, which are explained below.

3.3.2.1 Outcome 1: No learning

In this learning outcome, the acoustic values of two different L2 sounds (here, those corresponding to /ʌ/ and /ɑ/) are mapped onto the same perceptual L1 category (here, Spanish /a/) and are therefore perceived as the same sound. Considering that this mapping is best explained by the perceptual grammar of the L1, with only five possible candidates as targets for the mappings: /a/ and /o/, and thus following Schwartz & Sprouse's (1996) Full Transfer hypothesis, then this is in fact a no-learning scenario. This results in a situation where lexical items of a minimal pair (e.g. *cop* - *cup*) are activated by the same underlying representation (UR) /kap/, which was created during the initial state. The lack of a new perceptual category does not allow the L2 learner to create two different URs, and therefore the apparent ambiguity in the signal is resolved only by lexical context.

Figure 3.1 shows this outcome as a process where the acoustic values of /a/, /o/, /ʌ/, and /ɑ/ (in that order, top-down) are mapped onto the L1 categories /a/ and /o/. While the native category /a/ takes the mean acoustic values of Spanish /a/, /ʌ/, and /ɑ/, the native category /o/ only takes the mean acoustic values corresponding to /o/. These values are taken from the Bradlow (1995) comparative study between English and Spanish vowels (see Table 3.2 below). It is worth noting that according to the data, the mean values for /a/ and /ʌ/ overlap: they are almost exactly the same; this will lead to an important difficulty in terms of modelling, which we will see below.

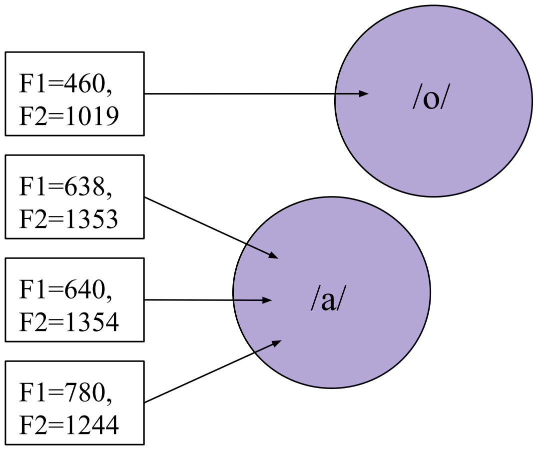


Figure 3.1: Outcome 1 - No learning. The acoustic values that correspond to /a/, /ʌ/, and /ɑ/, are being mapped onto the phonemic L1 category /a/.

Table 3.2: Acoustic values for vowels /a/ - /o/ - /ʌ/ - /ɑ/ (Bradlow, 1995)

	F1	**F2**
/a/	638	1353
/o/	460	1019
/ʌ/	640	1354
/ɑ/	780	1244

.

In order to formally define L2 speech perception processes, an OT approach has been proposed in the L2LP model by Escudero (2005), where a set of perceptual constraints define the way in which L2 sounds will be perceived by the listener. This framework takes into account a set of cue constraints that map the acoustic values of the input onto abstract phonological representations, specifically segments[6]. The following set of cue constraints would allow L2 listeners to categorise the sounds according to the categories that are present in their L1 phonological system (Spanish in this case). They are modelled after the set of constraints proposed in the L2LP framework, and taking into account the values in Bradlow's study.

- 640 Hz not /o/: an input of 640 Hz in F1 should not be categorised as /o/.
- 1350 Hz not /o/: an input of 1350 Hz in F2 should not be categorised as /o/.
- 640 Hz not /a/: an input of 640 Hz in F1 should not be categorised as /a/.

[6]While this work also states that a model involving an auditory-to-feature mapping followed by a feature-to-segment mapping might be possible, it does not include such level of granularity as it complicates the model without adding necessary information (p. 48, footnote).

- 1350 Hz not /a/: an input of 1350 Hz in F2 should not be categorised as /a/.
- 780 Hz not /o/: an input of 780 Hz in F1 should not be categorised as /o/.
- 1240 Hz not /o/: an input of 1240 Hz in F2 should not be categorised as /o/.
- 780 Hz not /a/: an input of 780 Hz in F1 should not be categorised as /a/.
- 1240 Hz not /a/: an input of 1240 Hz in F2 should not be categorised as /a/.
- *NEW: Do not create a new category for acoustic values that can be mapped onto existing categories. This constraint can be broken down into the following cue constraints:
  - 640 Hz not /ʌ/: an input of of 640 Hz in F1 should not be categorised as /ʌ/.
  - 1350 Hz not /ʌ/: an input of of 1350 Hz in F2 should not be categorised as /ʌ/.
  - 780 Hz not /ɑ/: an input of of 780 Hz in F1 should not be categorised as /ɑ/.
  - 1240 Hz not /ɑ/: an input of of 1240 Hz in F2 should not be categorised as /ɑ/.

*NEW is essentially a shorthand for the constraints related to L2-like categories /ɑ/ and /ʌ/, which are only present in outcome 4 (full category split). This constraint (or set of constraints) is ranked higher in the perceptual grammar for outcomes 1 to 3, but not 4.

Table 3.3: Perception of /ʌ/ by L1 Spanish speakers of English in a no learning outcome.

[640, 1350]	*NEW	640 Hz not /o/	640 Hz not /a/	1350 Hz not /o/	1350 Hz not /a/
a. ☞ /a/			*		*
b. /o/		*!		*	
c. /ʌ/	*!				

Tableau 3.3 shows a perceptual grammar that has been fully transfered from the L1 (thus with a highly ranked *NEW constraint), with the acoustic values for /ʌ/ as input. The constraints that do not allow for categorization of acoustic values into /o/ have been ranked higher, thus accepting those acoustic values as corresponding to exemplars of the /a/ category in Spanish. This perceptual grammar has no further candidates than those offered by the L1, which in this case are almost exactly the same as the mean values for /a/. If we are to assume that the mean acoustic values

of a certain sound can be considered category prototypes, then /ʌ/ is being perceived by a naive speaker as a perfect prototype of /a/.

Tableau 3.4 shows how the perceptual grammar of the initial/no learning state categorises the acoustic values corresponding to /ɑ/. Here the constraints against /o/ are again ranked lower, thus allowing for categorization of [F1=780 Hz, F2=1240 Hz] as /a/. Since this perceptual grammar has no candidates that could take F1 values higher than 640 Hz, (that is, the only [+low] vowel is /a/, which already takes F1=640 hz as [+low]), then /a/ is the only possible candidate. However, would sounds with the acoustic values for /ɑ/ be considered less prototypical exemplars of /a/ than tokens with the acoustic values for /ʌ/? This is a question that will be addressed more thoroughly in Chapter 5, but for the time being we will assume that the values for both /ʌ/ and /ɑ/ are perceived as /a/ with the same likelihood.

Table 3.4: Perception of /ɑ/ by L1 Spanish speakers of English in a no learning outcome.

[780, 1240]	*New	780 Hz not /o/	780 Hz not /a/	1240 Hz not /o/	1240 Hz not /a/
a. ☞ /a/			*		*
b. /o/		*!		*	
c. /ɑ/	*!				

It is important to note, however, that this is only a perceptual account, as L2 speakers may have different categories in production. In this respect, it is worth considering the claim made by Bohn and Flege (1997), according to whom even though experience in the L2 plays an important role in both production and perception of the English /æ/ - /ɛ/ contrast in native speakers of German, the latter is considerably more target-like than the former. Further experiments have also concluded that even though there is a correlation between production and perception, an even more important correlation is given by the L2 speaker's L1 as a predictor of difficulty in both producing and perceiving contrasts: The same /æ/ - /ɛ/ contrast is more elusive for native speakers of German (reglardless of their experience in the L2) than for native speakers of Spanish, given that native speakers of Spanish map these two vowels onto two different L1 categories. Conversely, native speakers of Spanish show less accuracy than native speakers of German in both production and perception of the /i/ - /ɪ/ contrast, given that the speakers/listeners of the latter language map these vowels onto two different L1 categories, while the speakers/listeners of the former language map them onto the same L1 category /i/ (Flege, Bohn, & Jang, 1997).

3.3.2.2 Outcome 2: Deflection

In a *deflection* scenario the learner has processed the acoustic values of the input from the environment and rearranged her perceptual grammar, but the learning outcome is not a category split and subsequent creation of two categories. Instead, one L2 category is attracted by one L1 category, and the other L2 category is attracted by a second L1 category, thus resulting in what Best (1995) defines as Two-Category (TC) assimilation. This would nevertheless mean that some learning takes place, but not in the expected way. Rather, the learner creates an L2 perceptual grammar that keeps the same L1 categories but maps the incoming acoustic signal into two different L1 categories. In this case, an initial mapping of the acoustic values corresponding to the English categories /ɑ/ and /ʌ/ onto the native category /a/ becomes a mapping of the values of /ɑ/ onto /a/ and those of /ʌ/ onto /o/. Figure 3.2 shows this outcome.

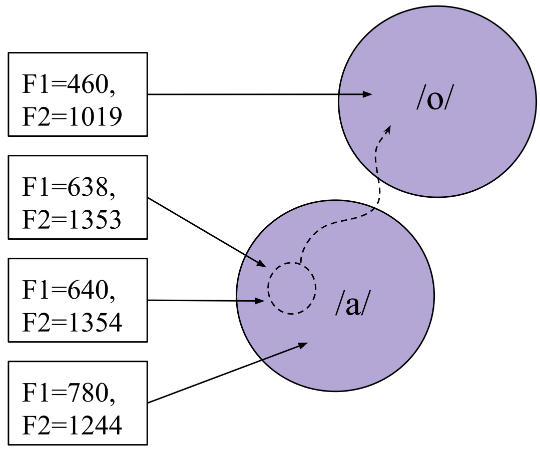


Figure 3.2: Outcome 2 - Deflection: tokens of one of the L2 categories (here /ʌ/) migrate to the neighbouring L1 category /o/ after having been initially mapped onto the L1 category /a/.

Although remapping seems to be a logically sound way to deal with the incoming input, two points must be taken into account. Firstly, such perceptual learning implies that /ɑ/ and /ʌ/ are first perceived as the same category, and *then* considered to be different, therefore getting a different mapping. In this regard, the reasons for such a change could be found in the development of a different cue-weighting strategy, which could be acquired as a result of large amounts of input and therefore a higher sensitivity to phonetic detail. And secondly, such a remapping of the acoustic signal could have two different outcomes: one where the acoustic values for /ʌ/ are mapped onto /a/ and /ɑ/ onto /o/ (Case 1) and another where the acoustic values for /ɑ/ are mapped onto /a/ and /ʌ/ onto /o/ (Case 2). However, and considering the data from Bradlow (1995), the values for /a/ are almost exactly the same as those for /ʌ/, which in Case 1 would make the values for /a/ also migrate to /o/; on the other hand, /ɑ/ retracts further back and lower in the perceptual space, thus making /a/ a better candidate if a more feature-based approach is taken.

Let us first deal with Case 1. While /ʌ/ would naturally map onto /a/ given that its production values are almost the same, why (and how) could the L2 listener categorise /ɑ/ onto /o/? The only way to make this possible is by assuming that an F2 of 1244 Hz is *too far back* to be a possible value for /a/, which would lead us to propose that perception in this case is mediated by a notion of prototypes based on features, and not based exclusively on acoustic distance. In order to model this outcome, constraints such as those proposed above can be rearranged in such a way that F2 values of 1240 Hz can be mapped onto /o/ . Tables 3.5 and 3.6 show this perceptual outcome; *NEW is again highly ranked.

Table 3.5: Deflection, Case 1: (a) acoustic values corresponding to /ʌ/ are mapped onto /a/.

[640, 1350]	*NEW	1350 Hz not /o/	1350 Hz not /a/	640 Hz not /o/	640 Hz not /a/
a. ☞ /a/			*		*
b. /o/		*!		*	
c. /ʌ/	*!				

Table 3.6: Deflection, Case 1: (b) acoustic values corresponding to /ɑ/ are mapped onto /o/.

[780, 1240]	*NEW	1240 Hz not /a/	1240 Hz not /o/	780 Hz not /a/	780 Hz not /o/
a. /a/		*!		*	
b. ☞ /o/			*		*
c. /ɑ/	*!				

On the other hand, Case 2 looks even less optimal if we assume an acoustic distance approach: /ʌ/ would have to leave a rather faithful mapping onto /a/ in favour of a category that is further away in the acoustic space. Tables 3.7 and 3.8 show how a constraint ranking similar to that of Outcome 1 (i.e. with height ranked higher than backness, but assuming a lower boundary) would lead to Case 2.

The deflection outcome has two implications that are worth noting. The first implication is that there is no creation of new perceptual categories; the perceptual representations of L2 sounds are being copied to the L2 grammar and are kept as such. The second implication is that the perception of the acoustic input would first map onto a phonological feature, and then to the category; this allows for not only perceptual warpings with high distorsions such as those proposed by this perceptual outcome, but also for changes in weighting of the different perceptual cues separately.

Table 3.7: Deflection, Case 2 (a): acoustic values corresponding to /ʌ/ are categorised as /o/.

[640, 1350]	*New	640 Hz not /a/	640 Hz not /o/	1350 Hz not /a/	1350 Hz not /o/
a. /a/		∗!		∗	
b. ☞ /o/			∗		∗
c. /ʌ/	∗!				

Table 3.8: Deflection, Case 2 (b): acoustic values of /ɑ/ mapped onto /a/.

[780, 1240]	*New	780 Hz not /o/	780 Hz not /a/	1240 Hz not /o/	1240 Hz not /a/
a. ☞ /a/			∗		∗
b. /o/		∗!		∗	
c. /ɑ/	∗!				

3.3.2.3 Outcome 3: Subsetting

This outcome assumes the same initial state as in the cases above: An SC assimilation type in naïve speakers, whose perception should progress to the learning scenario described by Escudero as learning a *new* sound. However, in Outcome 3 the learner achieves a different final state where instead of a full category split, two subsets of a same phonemic category are created. Figure 3.2 shows this outcome, where the difference in colours represent the different phonemic inventories, with L2-originated sub-categories in yellow; L1 categories (in purple) are not active. This poses a major problem in terms of phonological theory, as it goes against the notion of the phoneme as a contrastive unit; rather, such a learning outcome would create two different versions of what is labelled as the same sound.

The learning outcome from a representational point of view is a partial split of the category, where the L2 learner perceptually recognises the two sounds as members of the same L1 category, but can perceive them as somewhat different, particularly when the two contrasting L2 sounds are presented close in time to each other. The impression of the L2 learner is that these two different sounds in the L2 are “a different kind” of the L1 category.

Such a distinction could be aided by prototypicality: If one of the L2 categories is perceived as a better instance than the other (i.e. /ʌ/ is deemed a better instance of /a/ than /ɑ/), then discrimination of these two vowels should be easier. This has been described by PAM as Category-Goodness assimilation (Best, 1995, p. 195), which although formulated as an initial state scenario, it could also be seen as a potential learning outcome.

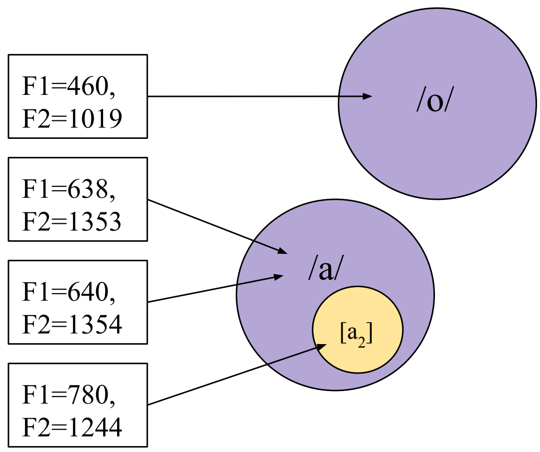


Figure 3.3: Outcome 3 - Subsetting: sounds with acoustic values for the L2 categories /ɑ/ and /ʌ/ are being mapped onto the L1 category /a/. The L2 speaker perceives /ɑ/ as different from prototypical values of /a/, but she still identifies them as a token of this L1 category.

At this point, the L2 speaker has reached a level where *noticing* has already taken place. From this point onwards, one might ask: what is stopping these learners from acquiring a new phonemic category? Firstly, an L2 learner that has acquired the language at a later stage without phonetic training has, as mentioned earlier, no reason to suspect that the vowel inventories for both languages are different. And secondly, as PAM predicts, two sounds that are perceived as exemplars of the same L1 category are not easily discriminated as different: even in a Category-Goodness assimilation scenario, discrimination is expected to be moderate at best. Here is where the constraint *New becomes relevant: the learner has noticed that something is different, but not enough so as to identify them as different phonemic categories.

However, the formalisation of this outcome presents a problem. A perceptual OT approach such as the one given in L2LP is unable to show a distinction between categories with different status in the inventory: since this is a situation where listeners would be able to perceive a difference only in a discrimination task but will assign the same phonemic label to them in an identification task, it cannot be considered a phonemic category. Unfortunately, a formalisation of this stage that assumes only phonemic categories as possible perceptual targets will yield the exact same output as that of a no learning outcome: the acoustic values for /ɑ/ and /ʌ/ mapped onto Spanish /a/ (See Tableaux 3.3 and 3.4).

Nevertheless, perception of acoustic details in terms of their closeness to other acoustic values can be modelled with a *Warp constraint, which ensures that a certain acoustic value will be mapped onto the closest category, *provided that there is one*. In this regard, Boersma (1999) states that "if *Warp is ranked higher than *Categ(other) for distances over 100 Hz, an input of [600 Hz] will be perceived as

/600 Hz/, perhaps creating a new weak category /lower mid vowel/[7]" (p. 4).

It is worth noting that here I depart from the L2LP model in terms of the formalisation: *Warp has not been taken into account in L2LP, since this constraint does not refer to language-specific phonemic categories. I present a solution where I assume the less prototypical subset sound as a phonetic category (with square brackets) but that is nevertheless phonemically the same; it represents a phonetic category delimited by a deviant, yet consistent over time, distribution in the input of what is perceived as a less prototypical instance of a phonemic category.

Tableaux 3.9 and 3.10 show this subsetting outcome, assuming that and [a_2] is a subcategory created after listening to the acoustic input corresponding both to /ʌ/ and /ɑ/, where [a_2] is a less prototypical value for the Spanish category /a/). I will assume for the sake of illustration that the acoustic values corresponding to /ʌ/ are considered by the listener as prototypical for the Spanish category /a/; this assumption is backed by the acoustic values found both in the study by Bradlow and this research. This category will be named simply /a/. On the other hand, the acoustic values for English /ɑ/ will be evaluated by the listener as a more deviant exemplar of /a/, which will in turn lead to the creation of the phonetic category [a_2]. The reason why I have employed the *Warp constraint for modelling creation of this phonetic category is just based on an acoustic similarity criterion, a point that will be discussed more thoroughly in Chapters 5 and 6.

Again, this learning scenario has forced me to depart from the convention proposed in Boersma (1999) regarding constraints. While *Warp has been originally formulated as an auditory-to-auditory constraint (e.g. *Warp([530], /740/), where the acoustic is given in square brackets and the target value is in slashes, thus standing for a phonological, yet not phonemic, category), it is somewhat inconsistent with regard to the rest of the cue constraints proposed here. Hence, Tableau 3.10 shows this as *Warp ([780], /a/), which can be defined as "Do not categorise an acoustic value of F1 = 780 Hz into /a/, whose prototypical acoustic value is F1 = 640 Hz".

Table 3.9: Perception of /ʌ/ in the subsetting outcome.

[640, 1350]	*New	640 Hz not /o/	640 Hz not /a/	1350 Hz not /o/	1350 Hz not /a/
a. ☞ /a/			*!		*
b. /o/		*!		*	
c. /ʌ/	*!				

In Tableau 3.10, the noticing process triggers a *Warp constraint that prevents a value of 780 Hz to be warped to the acoustic values for /a/. However, the *New

[7]Assuming that the only perceptual classes available are /260 Hz/, /470 Hz/, and /740 Hz/; namely, high, mid, and low vowels, respectively.

Table 3.10: Perception of /ɑ/ in the subsetting outcome.

[780, 1240]			*NEW	780 Hz not /o/	*WARP ([780], /a/)	780 Hz not /a/
a.		/a/			*!	*
b.	☞	[a₂]				*
c.		/o/		*!		
d.		/ɑ/	*!			

constraint is still highly ranked, thus blocking the creation of a new phonemic category to which the acoustic values corresponding to /ɑ/ could be mapped.

3.3.2.4 Outcome 4: Full category split

In a full category split, the resulting representations are different from those of the L1 and correspond to the set of representations that native speakers of the TL have in their inventory. Figure 3.4 shows this outcome, with the newly created L2 categories in yellow. They are available whenever the input is in the L2; the L1 categories remain inactive at that moment. Likewise, when the input corresponds to the L1 then the L2 categories are inactive. This conceptualisation is in line with the notion of language mode (Grosjean, 2001) mentioned in section 3.2.3, and it allows us to assume that the newly created categories will not activate when being in the L1 mode, and *vice versa*.

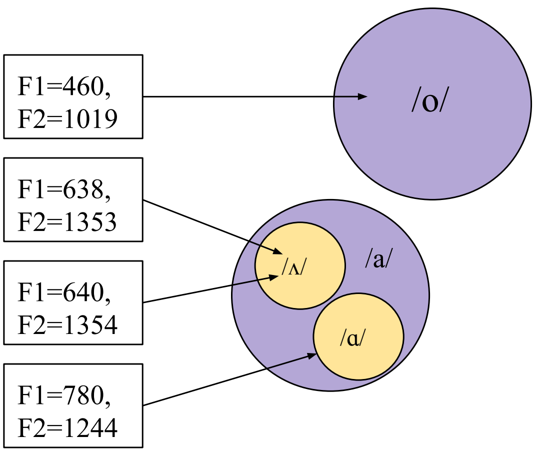


Figure 3.4: Outcome 4 - Full category split: tokens of the L2 categories /ɑ/ and /ʌ/ as new independent categories.

In Escudero's model it is assumed that constant input triggers a noticing process, where the L2 learner obtains feedback from the lexical context in order to disambiguate and choose the right lexical item; eventually, the different lexical representations will acquire the corresponding target-like phonemic representation. Despite being seemingly an acoustically-driven process, the evaluation of the learner's representation versus the actual input form plays a major role in the creation of a new perceptual category. Such a process has an important implication: That the L2 learner is at some point during the learning process, actually able to perceive acoustic detail in such a way that it allows them to notice the mismatch between the UR as formed by means of L1 phonological transfer (i.e. the form /kap/ and the input (i.e. either [kʌp] or [kɑp]), so that the creation of new URs /kʌp/ and /kɑp/ takes place.

A full category split implies a qualitative leap for the L2 learner. Whilst in an initial state the learner is unaware of the fact that the L2 has more vowels than the L1, a final state of category split would necessarily entail a process of *noticing* that what they thought were tokens of the same vowels as being slightly different, just like in the subsetting outcome. However, this is now being followed by a process of assigning to one of the vowels a different label, which allows the L2 speaker to clearly distinguish them in a minimal pair.

In order to model perception after the new category has been created, the constraint *New needs to be demoted in favour of a constraint that allows categorisation of the acoustic input as an instance of the new category, which ensures that the acoustic input is as close as possible to the center values for the existing categories. Tables 3.11 and 3.12 show the process; the constraint *New now appears broken down into the relevant cue constraints.

Table 3.11: Perception of /ʌ/ in a full category split.

[640, 1350]	640 Hz not /o/	640 Hz not /a/	640 Hz not /ɑ/	640 Hz not /ʌ/
a. /a/		*!		
b. /o/	*!			
c. ☞ /ʌ/				*
d. /ɑ/			*!	

Table 3.12: Perception of /ɑ/ in a full category split.

[780, 1244]	780 Hz not /o/	780 Hz not /a/	780 Hz not /ʌ/	780 Hz not /ɑ/
a. /a/		*!		
b. /o/	*!			
c. /ʌ/			*!	
d. ☞ /ɑ/				

However, the formalisation cannot show the reason why an input of 640 Hz is no longer recognised by the perceptual grammar as an instance of /a/, given that the prototypical acoustic values for both /ʌ/ or /a/ are the same; in this case, one must assume that the language mode will define whether categorisation of such input as either /ʌ/ or /a/ should take place. In this sense, and taking into consideration the assumption that categories that do not belong to the language being spoken at the moment are inactive, it could also be assumed that /a/ and /o/ are no longer possible candidates at this stage of acquisition.

Likewise, assigning a new label to a vowel sound that was originally mapped onto one of the L1 categories (shown as candidates c. and d. in Tables 3.11 and 3.12) is rather difficult from a cognitive point of view; particularly for L2 speakers whose L1 has transparent orthographic systems. L2 learners need to bootstrap themselves from a cognitive state where the incoming tokens in the input are perceived as an L1 category, a perception that is reinforced by the transparency of orthographic L1 knowledge: For Spanish speakers, *one letter equals one vowel*. Hence, the L2 learner does not know that the tokens being perceived as exemplars of category A are, in fact, members of categories X and Z in the TL; but more importantly, she has no way to know. This is particularly the case for L2 speakers who have gained enough proficiency to approximate their pronunciation to the target form, as NS will not notice anything different.

Furthermore, the L2 learner needs to create a target value for new vowels that do not have the same acoustic values in the L1 (in this case, we are assuming that /ɑ/ does not match any of the existing representations in Spanish). In order for this to happen, the L2 learner will need to perceive the difference between the values of the L2 category that matches an L1 vowel prototype (i.e. /ʌ/ and /a/), and those of the vowel that has no perfect match with any other vowel (namely /ɑ/), which -as previously discussed- are being perceived as different instances of the same vowel. This is somehow explained by the formalisation above, although it does not exactly explain how L2 speakers break the initial-state categorisation process, which *overrides* perception of the raw acoustic values in the input.

Additionally, the target values need to be stored as long-term memory representations, which should be retrievable online in order to compare them with *every instance* of the incoming input so that the correct URs be created. However, this cannot happen without the previous steps, for which we assume that a bottom-up approach to L2 category creation is, if not highly unlikely, not possible.

In sum, it is unlikely that this outcome takes place, mainly for two reasons. Firstly, it will be necessary to notice the difference between L1 and L2 categories, which is already unlikely if we take into account that L1 categories will warp the perceptual space, therefore making the discrimination of similar L2 vowels not easily accessible. And secondly, this does not seem to be possible without a process of declarative knowledge regarding differences in vowel inventories. Hence, it seems that the only way to achieve optimal L2 perception is by assuming an approach where L2 speakers are (a) aware of the difference between the new L2 categories, and (b) that they are able to convert that into *intake*, which entails processing the differences between the current state of knowledge in the L2 and the actual input form, and convert the new difference into newly acquired L2 knowledge (Ellis, 1994).

3.4 Towards a methodology for L2 category creation research

While the previous section has outlined the current theories and a few hypotheses regarding creation of categories in L2 speech perception (thus setting aside methodological concerns for a moment), here I turn to discuss methodological approaches, or how experimental research informs on the nature of perceptual representations in L2 learners. Perception requires an extremely careful planning of experimental designs in order to obtain the correct type of information.

What is native-like L2 speech perception? Naturally, the answer to this question would play a decisive role in choosing the right experimental approach. First of all, an important distinction must be made between category creation and speech perception. In this regard, Holt and Lotto (2010) point out that speech perception is a categorisation task, provided that perceptual categories are already there. This is exactly what a pseudo-homophony situation such as that of *cup-cop* in L1 Spanish speakers of English entails. A key assumption that we take is that category creation *precedes* speech perception, insofar as we understand perception (at least at this low-level stage) as a process involving categorisation: raw perceptual input needs to be recognised as a member of a category, be it phonological, lexical, or else, in order to be recognised as a phoneme, which will in turn form part of a word. With this in mind, it is worth noting that this work does not only look into the way in which L1 Spanish speakers of English categorise the sounds of the L2; it is also looking for evidence on whether hey create new categories, and how they could be described from a phonological perspective.

Each one of the wide range of perceptual tasks relates to different aspects of perception, which at the same time relate to different aspects of phonology. This brief review will attempt to summarise the different approaches to phonological issues that have been addressed from a perceptual viewpoint.

3.4.1 Frequently used methods

3.4.1.1 Identification

Identification (or categorisation) is perhaps one of the most common perceptual tasks in phonology. The subject is exposed to a single stimulus and is asked to give a certain response to the question: "What sound did you hear"? The usual procedure involves a set of possible predetermined responses, from which the subject has to choose one (a *forced-choice task*). Other possibilities are open-choice, where the subject can freely elicit a response after having heard the stimulus, or multiple-choice, where the subject can choose more than one option among a series of possible answers. (Colantoni, Steele, & Escudero, 2015). Responses can be represented orthographically, or through an image that either pictures a lexical item or contains a specific sound in the lexical item depicted.

Identification tasks are related to the categorical status of a certain sound. When the task aims to find out common miscategorisations among sounds, a confusion matrix can summarise the results and show the most frequent errors, assuming that the task had clearly defined stimuli.

3.4.1.2 Discrimination

Discrimination has usually been considered to be a less phonological task, and more related to phonetic aspects of perception (Colantoni et al., 2015). Discrimination tasks present more than one stimulus per trial. In an AX task, subjects are asked to answer whether two sounds are the same or different; an AXB task presents three stimuli, where two of them are the same, and the subject is asked to pair those that are considered to be the same. If the stimulus to be discriminated is after the reference stimuli, the experiment is called an ABX task. Conversely, in an XAB task the stimulus in question comes before the referential ones. All of these tasks have been considered to aim at different types of processing. It has been argued that AX tasks elicit acoustically-driven responses, while ABX, XAB, and AXB would be able to elicit phonologically-driven responses, due to the need for the listener to resort to phonological categories in order to group the stimuli (Colantoni et al., 2015, p. 97). However, it has been also proven that the order of the stimuli could also affect responses (Werker & Logan, 1985). It has also been argued that AX is a rather bias-free task, given that XAB and ABX are biased towards the closest stimulus, and given the cognitive load that implies storing two items in the short-term memory. On the other hand, is has also been argued that the interstimulus interval (ISI) and nothing else, is the crucial element in eliciting acoustic, phonetic, or phonological responses: The longer the ISI, the more likely it is that the response will come from the phonology domain (Colantoni et al., 2015, p. 97).

3.4.1.3 Rankings

Ranking tasks have different purposes. They return a continuous range of values that can be easily assimilable to a continuous variable, which can be particularly useful for phonetic research. However, if the research question is related to categoricity, then this task might not be the most suitable one. Nevertheless, it has also been proven to give interesting insights regarding questions related to the perception of nonnative sounds.

- Dissimilarity

  Dissimilarity is a task that aims to elicit responses related to a pair of sounds. An example of this type of task in L2 speech perception can be found in Flege et al. (1994), where subjects were asked to rate the degree of dissimilarity between two native vowels, two nonnative vowels, and one native and one nonnative vowel. Subjects were asked to rate the vowel pairs on a scale of 1 to 9, where 1 = similar, and 9= very dissimilar. The results showed that nonnative speakers of English (NL=Spanish), regardless of their experience in the L2, would rate vowel pairs similarly, assigning higher values to vowels that were considerably far in the acoustic space (such as /ɑ - i/) and a very low score for nonnative vowels that were adjacent in the acoustic space (such as æ- ɛ/. This methodology poses some problems for phonological theory, mainly that responses do not necessarily elicit any information with regard to the phonological status of a certain sound. However, it does give valuable information regarding the overall perception of differences along the entire perceptual space, as it can show perceptual assimilation of nonnative categories when ratings are too low, and categorical perception when they are high, without having to resort to a forced-choice test with category labels that nonnative speakers might not be able to recognise.

- Prototypes

  Prototype theory in speech perception was first proposed by Kuhl and Iverson (1995) and the Perceptual Magnet Effect described in Section 2.4.2.3. The theoretical claim that prototype theory could also apply to speech perception was tested empirically through a ranking task. In one of the experiments (Kuhl, 1991), subjects were asked to listen to a set of stimuli corresponding to synthesised tokens of the /i/ vowel. The stimuli were created from a vowel that was previously rated by subjects as a good instance (a prototype) of the vowel /i/, and another vowel sound that was rated as a poor instance (non-prototype) of the same vowel. Kuhl & Iverson synthesised the set of stimuli from these two reference vowels by altering the formant values in a radial manner, and asked their subjects to rate the tokens using a 1-to-7 scale, where 1= poor exemplar, and 7= good exemplar. The results showed that good ratings were given to the stimuli surrounding the prototype, and bad scores to the non-prototype surrounding stimuli. These results were used for a second experiment, where subjects had to discriminate between the reference sound (either the prototype or non-prototype) and one of their surrounding stimuli, in order to measure the sensitivity of subjects to the differences between acoustically close stimuli. In this case, rating were used not only to build the stimuli for a second experiment; the resulting stimuli from the manipulation of the prototype and non-prototype were also rated. The rating test has since then become frequent in perceptual tests, in order to validate prototypes of stimuli.

3.4.1.4 Oddity

Oddity tests can be seen as a variant of AXB tests, as they are built under the same principle (discriminate among three elements). Oddity tests also have two (or more) same stimuli and a deviant one, but in this case the subject is asked to identify the odd element instead of grouping the same ones. This type of experiment is aimed at eliciting responses regarding phonological categorisation. Flege et al. (1994) used this paradigm for assessing the categorisation of adjacent versus non-adjacent English vowels in native speakers of English and in nonnative speakers of English whose L1 was Spanish. The results showed that the non-adjacency condition yielded similar results for both groups of speakers, but not in the adjacent condition.

3.4.1.5 Categorical perception

It is often assumed that any response obtained in perceptual tasks that are performed along an X-step continuum between two categories showing the existence of a boundary is categorical perception. However, this is not the case. CP can be understood as a phenomenon and a method (see section 3.4.1.5, where the latter refers to the use of two different tasks (categorisation and discrimination) over the same stimuli, and then calculating the extent to which they correlate; it is expected that discrimination rises at the category boundary. However, such an approach has been mostly used in L1 perceptual experiments, given the need to have robustly formed categories and therefore an unequivocal way to categorise with labels.

3.4.1.6 Sensitivity

Detection theory is "a general psychophysical approach to measuring performance" (Macmillan & Creelman, 1991, p. 1), and more specifically, its goal is to measure accuracy in decision processes. It is particularly useful in speech perception, as categorisation and discrimination tasks are a way to measure the subject's performance usually under the assumption that there are correct and incorrect responses. In *correspondence experiments*, the subject needs to assign a certain response to different stimuli, provided that they have been paired with a specific correct response. The amount of discrepancy between the correct and the actual responses is what allows the assessment of performance. Errors can be divided into *Misses (M)* and *False Alarms (FA)*, and are attributed to *noise* that may come from either the subject or the stimuli in question. Sensitivity (*d'*) is the measure of the correct responses or *Hits (H)* minus the *FA* converted to a Z-score[8]. While it is very rare that subjects have a perfect score of 1, the H rate is usually greater than the *FA* rate (Macmillan & Creelman, 1991).

In an AX task, subjects are shown two stimuli per trial, and they are expected to decide whether these two stimuli are the same or different. Here it is expected that *d'* measures will be different for subjects who are native speakers of English than those who are not, as NS will have perceptual representations that NNS might not have. I will assume that more robustness in the same/different decision process will be due to the presence of different phonemic representations, rather than purely phonetic or non- differing representations at all. Hence, NS will show higher sensitivity measures to the /ɑ-ʌ/ pair than NNS.

Finally, a discrimination task that aims to measure sensitivity across different groups of subjects (thus looking for what we could call 'internal noise', i.e. presence/absence of perceptual categories) can assess the subjects' performance in two different ways. One is to consider correct responses using an objective approach, which entails having both AA pairs and AB pairs, so that only 'same' responses for AA (or BB) pairs are correct. A more language-specific approach would assume that correct responses are what NS do, so if they judge a certain AB pair as 'same', then NNS are not expected to perceive a difference; therefore, NNS who judge that same AB pair as 'same' would be providing a correct response.

3.4.2 Method assessment

The previous subsection has listed a wide range of experimental approaches in speech perception research. However, none of these by themselves would be able to tap the right kind of information needed for the purposes of this research; however, a combined approach could shed light over the issue. Here I review the previous methods and evaluate the way in which they contribute to the research question.

3.4.2.1 Identification and phonemic categories

While identification tasks tap phonological knowledge and are certainly able to provide information on the existence of phonemic categories, they might not be suitable in cases where subjects are non-native speakers, as they do not necessarily have the relevant labels that are needed in order to perform the task, particularly if the experiment considers both native and nonnative speakers. If an identification experiment uses only L2 labels for the stimuli, then the results would only inform between a no learning outcome, and a full category split. Subsetting and deflection outcomes would go under the umbrella of no learning. Likewise, using L1 labels with non-native subjects would force the interpretation of the responses under the

[8]The Z-score is calculated in different ways according to the type of task; here we will limit ourselves to follow McMillan and Creelman's approach.

perspective of L1 category reuse, and would not inform on whether new L2 categories have been created.

3.4.2.2 Perception of L2 categories with discrimination

As mentioned above, is is worth noting that discrimination tasks might point to different levels of representation. While identification is considered a linguistic task given that it makes use of the phonemic knowledge involved, discrimination is said to be able to move along the acoustic, phonetic, and phonological domains of speech perception (Colantoni et al., 2015, p. 97). In addition, it has been argued that the interstimulus interval (ISI) could also be crucial in tapping the right domain: The longer the ISI, the more the subject relies on phonological knowledge. However, research on this issue has not been able to test this hypothesis correctly (Werker & Logan, 1985; Gerrits & Schouten, 2004).

However, discrimination has an important advantage over other experimental tasks: it does not require labels for phonemic categories. This is relevant to our research insofar as we do not expect experienced L2 speakers to have them, but could be able to discriminate between sounds that are nevertheless assumed to have the same label: While they might be able to discriminate between /ɑ/ and /ʌ/ in a fairly consistent manner, they might still not identify them as two different categories.

A further aspect of discrimination tasks considers the difference between stimuli that are expected to be recognised as different. The simplest approach is to use stimuli that can be deemed as prototypes of two different phonemic categories by native speakers of the TL, and create different trials with the same two stimuli: A-A, A-B, B-B, and B-A. However, this does not provide information on the manner in which L2 speakers deal with the incoming acoustic stimuli: Aspects such as cue-weighting, or boundary movement are not elicited by this task. An alternative approach is to create continua between the two stimuli, which could vary either on just one perceptual cue, or simultaneously. This would certainly inform on the aspects mentioned above; however, and given the case that an advanced L2 speaker *would* have phonemic labels for the two sounds, it would be impossible to distinguish them from speakers who do not have them.

Furthermore, discrimination along stimuli drawn from a continuum leads to one more methodological decision: The acoustic distance of the stimuli to be compared in a trial. Discrimination of adjacent tokens provides information on phonemic boundaries, insofar as we accept the claim by Liberman et al. regarding enhanced discrimination across phonemic boundaries. However, if the categories that non-native speakers have created are not phonemic, then again we will not be able to determine whether a subsetting outcome has taken place: Boundaries will not show in the absence of phonemic categories. However, if all subjects are given a prototype stimulus of a certain category and are asked to compare it against any point along a continuum, then it is more likely to find a) a sharp, well-defined boundary (in the shape of an S-curve) in NS; b) no boundaries in naïve or inexperienced NNS (in the form of a flat line); and c) either a shifted boundary or a less steep S-curve in more proficient speakers.

3.4.2.3 Categorical perception in L2

Most CP tasks in speech perception take into account the methodology stated by Liberman et al. (1957), who created a continuum ranging between /b/ and /g/ and then created identification and discrimination tasks. However, the focus of the task was different, as the goal was to determine whether identification *determined* discrimination, thus providing evidence for impoverished discrimination within the boundaries of a certain category. Naturally, subjects were all native speakers of English.

However, observing the relation between discrimination and identification and comparing the results across groups of native and nonnative speakers would provide information on a potential subsetting scenario. Moreover, the same acoustic distances could be processed in a different way by the different speaker groups, which is the reason why comparisons between NS and NNS are necessary. In fact, the effect found by Liberman et al. regarding increased discrimination on different sides of a phonemic boundary is essentially a language-specific phenomenon, given that it relies on the existing phonemic categories.

3.4.2.4 Rankings

While prototype ratings are a straightforward manner to determine the value range of the category and its boundaries, it would not be useful if the listener lacks the category in question. However, comparing prototype rankings between native and non-native speakers would not yield any useful results, as L2 speakers might be making use of their L1 categories in order to complete the task. However, L2 speakers could rate L2-like stimuli as members of their L1 categories, which could shed light over the way in which L1 categories are reused.

3.4.2.5 Oddity

Since oddity tasks are basically discrimination tasks, the same pros and cons apply to them. However, there is one more aspect that needs to be taken into account: the cognitive load is much higher, considering that the subject needs to evaluate three different stimuli and make a judgement as a group.

3.4.3 A proposal for IL category research

The discussion above has shown that none of the methods described would be capable to test whether L2 speakers have created new categories, and the nature of the category that has been created at the same time.

3.4.3.1 Experiment 1

The most immediate concern is whether non-native speakers are able to leave the initial state. Since an identification task would be too difficult for L2 learners, discrimination would be able to provide more fine-grained results. Experiment 1 aims to provide a comparison among three groups: Experienced L2 English learners with Spanish as L1, less experienced L2 learners, and native speakers of English. In this regard an in-between behaviour in discrimination in experienced L2 speakers is expected, assuming that (a) the group of less experienced L2 learners would show no discrimination between /ɑ/ and /ʌ/; and (b) that native speakers of English, on the other hand, are expected to show excellent discrimination.

In order to observe the emergence of a category boundary, continua should be built. While a /ɑ - ʌ/ continuum would provide evidence regarding these two categories, continua built between /ɑ/ and Spanish vowels /a/ and /o/ would show where /ɑ/ is being originally mapped onto; likewise, continua between /ʌ/ and Spanish /a/ and /o/ would show the initial mapping of /ʌ/. If nonnative speakers can discriminate between e.g. /ʌ/ and /o/, but not between /ʌ/ and /a/, then /ʌ/ has been perceptually assimilated to /a/. Furthermore, it would provide evidence regarding the initial state: If both English vowels are mapped onto the same vowel, then the initial state is clearly one of SC assimilation, as predicted by PAM.

3.4.3.2 Experiment 2

Should Experiment 1 yield results towards a more native-like perceptual pattern in experienced nonnative speakers, then one further step should be taken: To test for the presence of identifiable phonemic categories that can be invoked by using L2-like labels. This could be considered sufficient evidence for a full phonemic split; however, and given the list of difficulties that L2 speakers would need to overcome in order to achieve such an outcome, their ability to distinguish between /ʌ/ and /ɑ/ could be better observed through a different discrimination task. However, identification could provide clear evidence for a deflection outcome, in which case a clear change in categorisations along the continuum is expected; otherwise, labelling would be random.

Identification with L1-like labels should, on the other hand, provide the evidence that the L2 labelling cannot provide: Could it be the case that speakers just do not have the right labels to encode the distinction? In this regard, L1 labels could help insofar as they could show a boundary, or at least, something other than random behaviour.

Since it would not be impossible to find evidence for excellent discrimination but no ability to create new labels, a discrimination task that looks for a boundary effect *á la* Liberman et al. would yield similar behaviour in both nonnative and native speakers of English, though without the label. Hence, in order to obtain fine-grained data, a 1-step and 2-step discrimination task using the same stimuli in identification would be needed. If experienced nonnative speakers of English do have clear boundaries between the two sounds but no label, then the same discrimination peak that is expected to show in native speakers should appear. Sensitivity measures (*d'*) should provide a more accurate account of discrimination.

Finally, a prototypicality rating task would provide evidence towards the presence of a category, this time without having a reference sound stored in short-term memory as in discrimination tasks. If there is a category boundary, be it phonemic or not, the ratings along the continuum would show a broken line pattern: They would go from high to low, and then back to high. Otherwise, a decreasing line should be expected. On the other hand, ratings with L1-labels would provide evidence towards the way in which L2 categories overlap with those in the L1: How prototypical of /a/ is, for instance, a token of /ɑ/? Is it the same as /ʌ/? If so, then, and according to PAM, discrimination would be extremely poor; if not, then IL representations could be better described by PAM's Goodness-of-fit assimilation. Furthermore, if experienced L2 speakers show high performance in the discrimination tasks, it would be worth testing whether these potential differences in discrimination would correlate with how prototypical they perceive that /ɑ/ and /ʌ/ are with regard to /a/.

3.5 Summary

This chapter had two main goals. First, it addressed the issue of category creation in L2 learners. Three L2 phonology models that have somehow dealt with the process of category creation in L2 were described and analysed: the Speech Learning Model, The Perceptual Assimilation Model, and the L2 Linguistic Perception model. However, only SLM offers a hypothesis regarding the nature of these newly created categories, claiming that they are phonetic in nature; on the other hand, only L2LP offers a formal account. I have hypothesised around four possible learning outcomes: no learning, deflection, subsetting, and full category split, arguing also that the split is unlikely to occur. These learning outcomes have been modelled formally by following mainly L2LP in most aspects, but departing from it in others.

And secondly, the chapter discusses several experimental approaches to L2 speech perception and the way in which these inform on the research question. I have argued in favour of a mixed approach including several tasks over the same set of stimuli, which should be able to provide information on different aspects of L2 speech perception.

CHAPTER

FOUR

LEAVING THE INITIAL STATE: PERCEIVING DIFFERENCES ACROSS THE PERCEPTUAL SPACE

Usté no es ná,
ni chicha ni limoná.

Ni chicha ni limoná
Víctor Jara

4.1 Introduction

The previous chapters have outlined the issue of L2 category creation when the L1 phonological inventory has fewer elements than the TL inventory. In this case, during the initial stage of acquisition, an L1 category will receive tokens of two different L2 categories, a situation that PAM refers to as a Single-Category (SC) assimilation and that we have called a 2-to-1 mapping (section 2.6). This research looks for evidence with regard to an 'end state'[1], in which these two English vowel categories could be perceived by highly proficient L1 Spanish speakers of English in the same (or similar) way as L1 English speakers do, i.e. ability to perceive the contrast between the two English vowels /ɑ/ and /ʌ/, which are perceptually mapped onto one L1 category /a/ by L1 Spanish learners of English. Such perceptual learning entails a potential L1 category split, where a certain range of acoustic values that were originally mapped onto a certain L1 category will have to split into two different ranges, to which these two English vowels will be associated. Moreover, it is expected that the degree of accuracy showed by L1 Spanish speakers of English with a high command of their L2 will at some point become similar to that of L1 English speakers, in the sense that they come to use similar perceptual cue-weighting strategies and acquiring a similar system of perceptual representations.

This Chapter presents Experiment 1, which aims to obtain evidence of changes (or lack thereof) in perception as the learner gains proficiency; the possibility that

[1]For lack of a better term; more precisely, an IL state other than the initial state.

the L2 learner would not change her perceptual grammar was described in Chapter 3 as a no-learning outcome (section 3.3.2.1). In order to achieve this, data with sensitivity measurements between the vowels /ɑ/-/ʌ/ was collected. Likewise, the experiment tested whether the subjects can discriminate along an acoustic 7-step continuum between these two vowels according to the subject's amount of experience in the L2. The experiment consisted of an AX discrimination task, where subjects of three different groups (Native speakers of English, advanced L2 speakers of English, and beginner L2 speakers of English) gave same/different responses to trials with two stimuli taken from an acoustic vowel-to-vowel continuum.

4.2 Leaving the initial state

Let us remember that the initial state in L1 Spanish speakers of English can be best described by a state of full copying of the L1 vowel inventory and perceptual strategies over the blank slate of the L2. Hence, the initial state of L1 Spanish speakers in perception of the English /ɑ - ʌ/ contrast is that both vowels will be perceived as the native Spanish category /a/, which causes words with different URs in the TL such as /kʌp/ *'cup'* and /kɑp/ *'cop'* to be perceived as homophones. The acoustic values of both /ɑ/ and /ʌ/ will be mapped onto the perceptual area containing the range of possible acoustic values for the category /a/, thus creating the same UR for both words: |kap|. As mentioned in Chapter 3, it is possible that these L2 speakers of English may not move forward in their /a-ʌ/ acquisition process, which has been described as a **no learning** outcome; this is the least optimal of all the possible learning outcomes.

Why should one think of the possibility of a no learning outcome? Three reasons should be considered. Firstly, research on the Perceptual Magnet Effect (Kuhl & Iverson, 1995), has provided evidence according to which good exemplars of a certain sound category (as rated by native speakers) behave in a different way than poor exemplars. Good exemplars of a category act as "perceptual magnets" that *warp* the surrounding perceptual space, so that tokens of a certain sound category that are acoustically close to what a listener perceives as prototypical will then be perceived as the same sound. On the other hand, sounds that are considered to be non-prototypical do not have this effect on their surrounding tokens, even if they are acoustically close. This finding has implications for non-native speech perception, as L1 categories might be acting as prototypes: regardless of how much L2 listeners strive to maintain a contrast between non-native sounds that are acoustically similar to an L1 category, this L1 category will warp their perception so as to inhibit perception of the difference. Figure 4.1 shows the closeness of the tokens of both /ɑ/ and /ʌ/ to the native L1 category /a/, where the latter works as a perceptual magnet that warps perception of the former two vowels.

Secondly, the phenomenon of categorical perception (Liberman et al., 1957) predicts poor within-category discrimination; hence, if L1 Spanish learners of English perceive /ʌ/ and /ɑ/ as just one category, then again, the expected discrimination between these two vowels is poor.

Thirdly, the input-to-intake problem refers to the lack of negative evidence in speech perception and its consequences for further perceptual learning. In order for L2 learners to leave their initial state during their acquisition process, they would have to at least be able to *notice* the difference between the L2 input and their mental representations (which, if we are to assume a Full Copying initial state, will be conformed by segments of the L1 sound inventory). However, this is unlikely if we consider the implications of the phenomena mentioned above.

Nevertheless, the presence of the aforementioned phenomena should not be understood as a reason for inevitable stagnation in perceptual learning. In fact, experiments show that advanced L2 speakers move towards patterns similar to those of native speakers as they gain more experience in the L2 (MacKain, Best, & Strange, 1981; Bohn & Flege, 1997; Flege et al., 1994), which suggests that there is a moment in which the perceptual constraints are overcome in order for L2 speakers to perceive differences between these categories. Thus, a reasonable assumption is that when L2 speakers have fewer sound categories in their native inventory than those present in the language being learnt, then some type of reorganisation of their phonological knowledge takes place. However, even though it has been proven that L2 speakers can learn how to perceive L2 phonetic contrasts that are not present in the L1, such results *do not necessarily imply that a new category has been created.* Different tasks may yield different results, as discrimination of similar stimuli depends more on changes in processing strategies and not a possible sensorineural loss, which explain why perception of nonnative contrasts is available only in certain tasks (Werker & Tees, 1984).

Hence, a first approach to the problem of creating perceptual categories is to investigate whether there is a moment in which L1 Spanish speakers of English become able to notice a difference between two L2 vowels that are being perceptually mapped onto one L1 vowel. At this point, we are not necessarily aiming for answers regarding the nature of such learning; rather, we are looking into the possibility that input may be perceived as different, but without any subsequent intake. One way to reach this goal is by looking into sensitivity to minimal pairs in non-native beginners, non-native advanced, and native English speakers; a second option is to observe the totality of the acoustic space between the relevant tokens and look for boundaries in perception that account for gradual changes.

4.3 Hypotheses

1. Regarding the **initial state**, I hypothesise that L1 Spanish learners of English with little experience in the L2 will show a perceptual merge of /ʌ/ and /ɑ/, which can be observed in two different ways:

    - Lack of perceptual contrast between /ʌ/ and any of the tokens in the /ɑ - ʌ/ continuum; likewise, a lack of perceptual contrast between /ɑ/ and any of the tokens in the /ɑ - ʌ/ continuum is expected.
    - Low sensitivity (*d'*) scores when comparing the first and the last token of the /ɑ - ʌ/ continuum.
    - Lack of perceptual contrast between /ɑ/ and any of the tokens from the /ɑ - a/ continuum. Similarly, lack of perceptual contrast between between /ʌ/ and any of the tokens in the /ʌ - a/ continuum is predicted.

2. Regarding the **IL state of experienced L1 Spanish learners of English**, I hypothesise that they will be able to perceive the difference between /ɑ/ and /ʌ/ though not with the same accuracy as native speakers. This could be observed as:

    - Probabilistic discrimination between /ʌ/ and the first half of the tokens in the /ɑ - ʌ/ continuum, with increasing 'different' counts at some point along the continuum. The same behaviour is expected between /ɑ/ and the second half of the tokens in the /ɑ - ʌ/ continuum.
    - Medium sensitivity (*d'*) scores when comparing the first and the last token of the /ɑ - ʌ/ continuum.
    - Lack of perceptual contrast between /ɑ/ and any of the tokens from the /ɑ - a/ continuum. Lack of perceptual contrast between /ʌ/ and any of the tokens in the /ʌ - a/ continuum is also predicted.

3. Regarding **native speakers of English**, I predict presence of a full perceptual contrast between /ɑ/ and /ʌ/, which can be observed thus:

    - Ceiling discrimination between /ʌ/ and tokens of the first half of the /ɑ - ʌ/ continuum, with an abrupt decrease in 'different' counts after the boundary (presumably token 4) is reached; an S-shaped curve is expected. The same behaviour is expected between /ɑ/ and tokens from the second half in the /ɑ - ʌ/ continuum.
    - High sensitivity (*d'*) scores when comparing the first and the last token of the /ɑ - ʌ/ continuum.

4.4 Methodology

4.4.1 Stimuli

Five 7-step vowel continua were created: one between the English vowels /ɑ/ and/ʌ/, and four between each one of the former and the Spanish vowels /a/ and /o/. Stimuli were built with Praat (Boersma & Weenink, 2012) from recordings of real tokens by female native speakers of General American English (GAE) and Chilean Spanish (ChS). The GAE vowels were obtained by asking the participant to elicit CVC words (*cup, cut, pot, top*) 3 times each in the carrier sentence *I say* _________ *once*. The ChS vowels were obtained by asking the participant to elicit the Spanish CVC words *pan, par, ron, por* in the carrier sentence *Yo digo* _________ *de nuevo*. The resulting continua were /ɑ-ʌ/, /ɑ-a/,/ɑ-o/, /ʌ-a/, and /ʌ-o/. Formant values of the original vowels were modified in equal Hertz[2] steps moving towards the vowel at the other endpoint, and were then re-synthesised using the same source. The final stimuli consisted of CVC nonce words where C is a voiceless plosive sound and V is any of the resulting continuum tokens, followed by an ISI of 1 second, and an isolated vowel corresponding to either one of the ends of the continuum. The resulting stimuli can be described as a vector, where AB is a continuum between sound A and sound B; and the subindices represent the continuum step. Since continua go from A to B and include 7 steps, AB_1 is A, and AB_7 is B. Equation (4.1) shows a stimulus where the CVC string has any of the possible resulting tokens from the continuum, followed by the ISI and the isolated vowel, which is the first endpoint of the continuum:

$$Stim = C[AB_i]C + ISI + A \tag{4.1}$$

Likewise, equation (4.2) shows a stimulus where the isolated vowel at the end is the other endpoint of the continuum. Note that the isolated vowels at the end are always continuum endpoints; this experiment did not consider stimuli where discrimination was performed over randomly drawn or adjacent token pairs.

$$Stim = C[AB_i]C + ISI + B \tag{4.2}$$

Table 4.1 shows the F1/ F2 differences between the endpoints of each continuum, and the corresponding Euclidean distance (ED). Vowel duration was equalised in order to avoid the use of duration as a perceptual cue. We will add, as a reference, the values between native Spanish vowels /a/ and /o/, even though perception of

[2]Instead of using a perceptual scale such as Mel or Bark, I decided to employ Hertz in an effort to establish a clear relationship between the production values for each category and the way in which these values are perceived. Likewise, since the main goal is to compare NNS perceptual patterns against those of NS, an experiment with stimuli varying according to a more precise logarithmic scale would not lead to different conclusions.

this contrast was not measured.

Table 4.1: Differences in Hz between endpoints of continua

Continuum	Δ F1 (Hz)	Δ F2 (Hz)	ED (Hz)
/ɑ - ʌ/	150	388	415
/ɑ - a/	29	252	253
/ɑ - o/	367	330	493
/ʌ - a/	121	136	182
/ʌ - o/	217	718	750
/a - o/	369	613	718

4.4.2 Subjects

The experiment considered 3 groups of subjects: a first group of L1 American English (NS, N=8) as a control group, a second group of native speakers of Spanish with some knowledge of English (NNS-B, N=13), and a third group of native Spanish speakers with advanced knowledge of English (NNS-A, N=9). The group of NS were all native speakers of American English. The NNS-A group were speakers of Chilean (8) and Mexican (1) Spanish who had lived in an English-speaking country for a minimum of 1 year and had passed the required IELTS/TOEFL exams to begin studies at University level. The NNS-B group were speakers of Chilean (12) and peninsular Spanish (1) who had lived in Manchester (UK) for up to 3 months.

4.4.3 Procedure

The experiment was carried out in the Phonetics laboratory at the University of Manchester, UK. For stimulus recording, both the American English and the Spanish speaker were recorded in a sound-attenuated booth using an AKG condenser microphone, model C520. The sounds were recorded on a HP ProBook 6570b with the Praat recording interface, at a sampling frequency of 44100 Hz.

Before the experiment, subjects were first asked to fill out a language background questionnaire (Annex 2). After completing the form, subjects were asked to perform an AX task, discriminating between the vowel in the nonce word and the one after the ISI. The experiment was set up in E-Prime, and subjects performed the task on computers in a sound attenuated booth using circumaural headphones. Stimuli were presented in random order and subjects had a maximum time of 4 s to respond. Subjects were asked to respond by pressing a green button whenever they thought the vowel sounds in the stimuli were the same, and a red button when they were different (the screen on the computer showed the words “Same” and “Different” in the corresponding colour and aligned with the buttons). For these purposes a

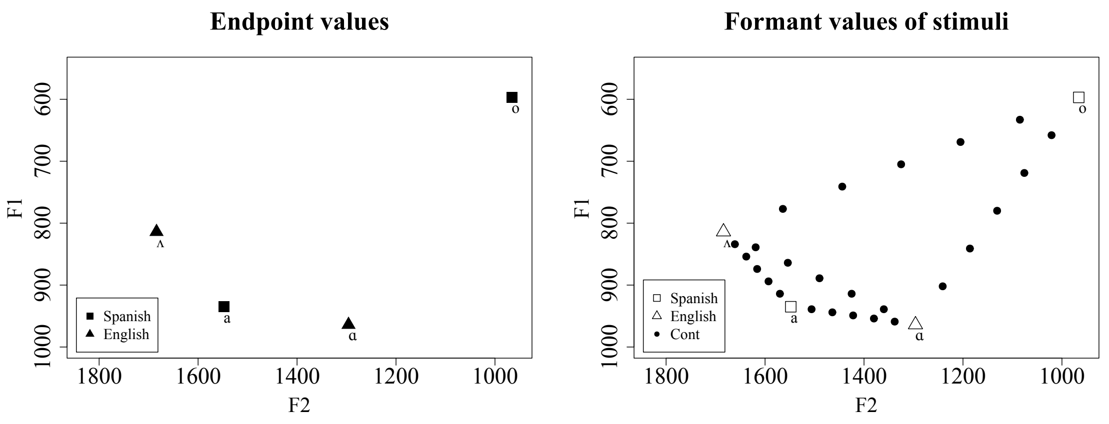


Figure 4.1: Formant values of endpoints and the resulting stimuli

Psychology Software Tools serial response box model 200a was used. The experiment had a total of 2 blocks with 35 trials each. All response times were recorded.

4.5 Results

4.5.1 Endpoint sensitivity

In order to obtain sensitivity scores, the results from two types of trials were analysed: (a) trials where the stimuli to be compared were the last token of the continuum, and the first token of the continuum after the ISI ($i = 7$ in equation 4.1); and (b) trials where the stimuli for comparison were the first token of the continuum, and the last token of the continuum after the ISI ($i = 1$ in equation 4.2).

Sensitivity was calculated with the procedure proposed by Macmillan and Creelman (1991) for same/different tasks with a differencing model, that is, for experimental designs with roving stimuli. In these cases the strategy used by the speaker is to assess the difference between two stimuli pairs according to an internal criterion, and thus not relying on previous observations throughout the experiment. The usual calculation for *d'* is given by the equation *d' = z(H) - z(F)*, where *z(H)* represents the z-values of the proportion of hits (i.e. the times when the subject clicked on "different" when the stimuli were different) and *z(F)* the z-values of the proportion of false alarms (that is, when the subject clicked on "different" when the stimuli were the same). However, the differencing strategy modifies the equation in order to represent the decision space according to a decision criterion given by the difference between two observations.

The *d'* scores were calculated using the *sensR* package for R (Christensen & Brockhoff, 2016), which includes the differencing strategy in the calculation and the corresponding standard errors (see Table C.1 in Annex 3). Given that only four

Table 4.2: Endpoint sensitivity (*d'*)

Continuum	NNS-B	NNS-A	NS
/ɑ - ʌ/	2.57	3.93	4.99
/ɑ - a/	0.88	1.87	2.43
/ɑ - o/	6.28	5.14	5.75
/ʌ - a/	0.87	0	2.43
/ʌ - o/	4.70	5.23	5.75

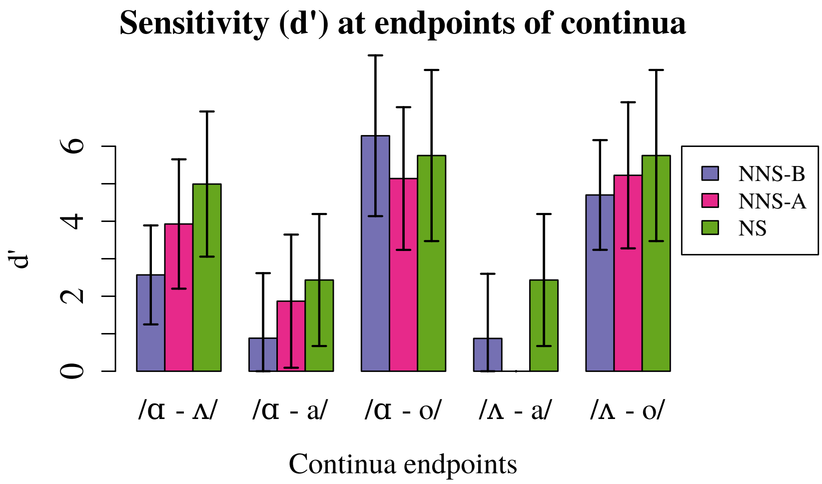


Figure 4.2: Endpoint sensitivity, with error bars showing a 95% confidence interval.

datapoints per subject were available (two same and two different stimuli pairs) the results were pooled. Additionally, observations from subjects who were too slow to provide a response (that is, they took longer than 4 s) were not taken into account in order to obtain the *z(H)-z(FA)* ratios. Since zero counts in H or FA would not allow to compute *d'* (as they yield infinite values) several correction measures have been proposed, which were summarised in Stanislaw and Todorov (1999); I have used the loglinear approach (where 0.5 is added to both H and FA, and then divided by N+1) to correct the H and FA rates of zero or one. The values given by the *sensR* package were obtained using a Thurstonian approach to sensitivity, with calculations for a same-different test further explained in Christensen and Brockhoff (2009); nevertheless, the values are exactly the same as those given by the Macmillan and Creelman method. Table 4.2 and Figure 4.2 show the scores obtained by each group for each continuum.

Each group showed different sensitivity (*d'*) measures, with both continuum endpoints and proficiency level in English as factors affecting the group scores. The *d'* scores among groups for the /ɑ/ - /ʌ/ distinction show an important difference. As expected, subjects from the NS group have higher sensitivity to prototypical tokens of /ɑ/ and /ʌ/ (*d'* = 4.99). On the other hand, albeit being the group with lowest *d'* score, the NNS-B group showed higher sensitivity than the expected (0 = no sensitivity), with *d'* = 2.57. Finally, the NNS-A group showed intermediate sensitivity (*d'* = 3.93).

The two vowel pairs that subjects were least sensitive to are /ɑ/ - /a/ and /ʌ/ - /a/; however, the NS group shows in both cases the highest *d'*. Interestingly, sensitivity measures in NNS-A are different for the two cases: while NNS-B and NS seem to be equally sensitive to both pairs, the NNS-A group shows an important change: from no sensitivity in the /ʌ/ - /a/ contrast, to low-moderate in the /ɑ/ - /a/ pair. This suggests that NNS-A become more sensitive to the front-back difference (which in this case is given by an F2 difference of 252 Hz; not an important difference for F2) than to the height contrast (with an F1 difference of 121 Hz; a much more important difference). Hence, even though /ɑ/ and /a/ are acoustically quite similar, NNS-A are more sensitive to the difference between this vowel pair than to /ʌ/ and /a/: for this group, /ɑ/ *may* be /a/, but /ʌ/ is *certainly* /a/.

On the other hand, the highest sensitivity was found in discrimination of the /ʌ/ - /o/ and /ɑ/ - /o/ contrasts. While the /ɑ/ - /o/ pair triggered a ceiling effect for all groups, /ʌ/ - /o/ did not: although still high, this pair triggered lower *d'* values in NNS-B, and an in-between value in NNS-A. NS also show the highest sensitivity levels. In this respect, it is clear that all subject groups have high sensitivity to the /ɑ - o/ pair, followed by /ʌ - o/. However, while the three groups showed to have a rather high sensitivity to the latter, they have different sensitivity to the former, with NNS-B speakers being two points below the perfect sensitivity showed by the NS group. This is the opposite situation: while /ʌ/ is very likely to not be /o/, /ɑ/ is certainly not /o/.

Regarding discrimination of native/nonnative pairs, two important remarks must be made. Firstly, there are two visible trends: while two endpoint pairs show high *d'* measures (namely, /ɑ/ - /o/ and /ʌ/ - /o/), the other two continua (/ɑ/ - /a/ and /ʌ/ - /a/) show much lower scores, hence pointing to the presence of noise in the stimuli, which can be attributed to the short acoustic distances. And second, that the NS group is *always* the group with higher sensitivity, which shows that both NNS groups are less sensitive to smaller formant frequency changes in the stimuli.

In sum, while all groups agree on that both /ɑ/ and /ʌ/ are perceptually closer to /a/ than to /o/, it is the degree of overall certainty that varies. The only contrast where all groups do not seem to agree is /ɑ/ - /ʌ/; here, both NNS groups have lower sensitivity than the NS group.

4.5.2 Changes along the continua

I will now move on from endpoint-to-endpoint sensitivity to continuum-to-endpoint behaviour, in search of boundary movements that illustrate how perception of contrasts takes place across the L2 acquisition process. AX results provided by subjects in the different continua show different group-specific patterns. While some continua show clear categorical effects with steep slopes in the S-curve, others show signs of gradient perception with gradual slopes, and others even show no signs of perception of differences at all, with flat lines near 0 in the Y-axis.

I will adopt the assumption that perceptual boundaries can only be found when the endpoints of a certain continuum are perceived as members of two different categories. If no boundaries are found, then the endpoints will be perceived by the listener as members of the same category. However, since the second element of the stimulus pair is always an endpoint, a boundary will not take the form of a peak, as when comparing adjacent continuum tokens. Rather, this will take the form of: (a) a flat horizontal line close to the X-axis, when no boundaries are present; (b) a raising or decreasing slope, when boundaries are not fixed; or (c) an S-curve, with an abrupt change in 'different' counts when the stimulus pair are on different sides of the boundary.

The Y-axis, on the other hand, shows the probability of 'different' responses given by a certain group. In this regard, the expected scenario in the NNS-B group for the /ɑ - ʌ/ continuum -who should be unable (or able to very little extent) to perceive the /ɑ/ - /ʌ/ contrast- is that their line of 'different' responses ought to lie close to 0 on the Y-axis. On the other hand, NS should present an abrupt rise from 0 to 100% 'different' responses, thus adopting an S-curve pattern.

However, none of the groups are expected to show perfect category boundaries in native-to-nonnative continua, unless they were perceiving the nonnative category entirely as another native category. For instance, and according to the sensitivity results presented above, this can be expected in the /ɑ - o/ continuum, but not in /ɑ - a/.

The data were modelled with a logistic regression, with the level of proficiency (i.e. group) and the first stimulus in the stimuli pair[3] as predictors for the response 'same' or 'different'. Models considering interactions and each variable as a single predictor were also tested in terms of their goodness of fit with an analysis of deviance. Given that endpoint-to endpoint discrimination gave sometimes perfect discrimination scores with values of 0 and 1, stimuli 4 was used as the reference level; regarding proficiency, group NNS-B was considered as reference. A full detail of the p-values and AIC obtained can be found in Annex 3. The error bars shown in the following graphs in this chapter correspond to a 95% confidence interval.

4.5.2.1 Perception of open back unrounded vowel /ɑ/

Two continua were built from the open back vowel /ɑ/: one moving towards /a/ (ED = 253 Hz) and another moving towards /o/ (ED = 493 Hz). The /ɑ - o/

[3]Stimuli pairings were modelled as a categorical variable, considering that interpolation data can only be valid when assuming linearity in responses along a scale; in this regard, perception along a Hertz scale will not be linear.

continuum showed a clear change in discrimination, moving from 100% counts of 'same' judgements to 100% 'different' judgements along the continuum in all groups (Figure 4.3). This can be regarded as evidence that all three groups perceive the endpoints as members of two different categories, and that this distinction leads to a sharp boundary in the middle of the continuum. As will be later discussed in section 4.6, the lack of significant differences in the curves suggests that the signal is perceived equally clear by all groups. The best-fitted model for the /ɑ - o/ versus /ɑ/ set of data considered only the stimuli pair as a predictor; proficiency level did not improve the model significantly. Discrimination against /o/ was also modelled with the first stimulus of the pair as a predictor: addition of proficiency level did not yield a significant addition to the fit. The fitted values were the same for all groups.

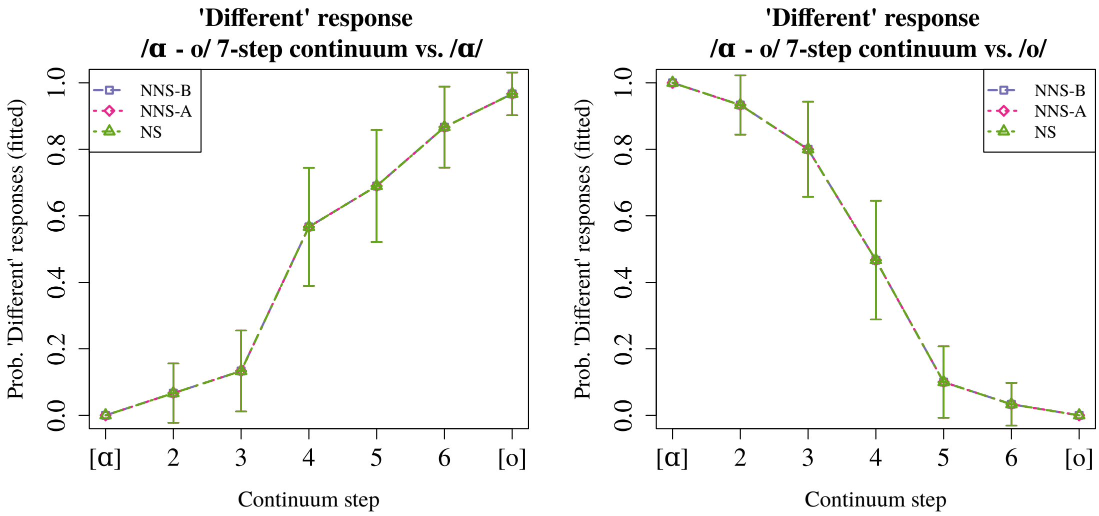


Figure 4.3: Results of AX test, /ɑ - o/

On the other hand, the /ɑ - a/ continuum showed very low discrimination between the stimuli present in each trial, where all groups showed a small peak when comparing /a/ to /ɑ/. This suggests that /ɑ/ and /a/ are perceived as members of the same category by subjects from not just the NNS groups but also by the NS group. The low discrimination values can be explained by the fact that both vowels /ɑ/ and /a/ are rather close in terms of their acoustic values.

Furthermore, the results given by the deviance analysis of the models suggest that there is little difference in perception of the acoustic differences along the /ɑ - a/ continuum in all groups (Figure 4.4); in fact, no model provided a good fit for the data as neither the stimuli pairs nor the proficiency level seemed to trigger a significantly higher rate of 'different' counts. Only a regression for the /ɑ - a/ - /a/ discrimination task with proficiency level as predictor gave a low significance level for the fit (p=0.0436).

It is also worth noting that these continuum steps are shorter than in the previous case, where the endpoints were farther from each other in the acoustic space. The

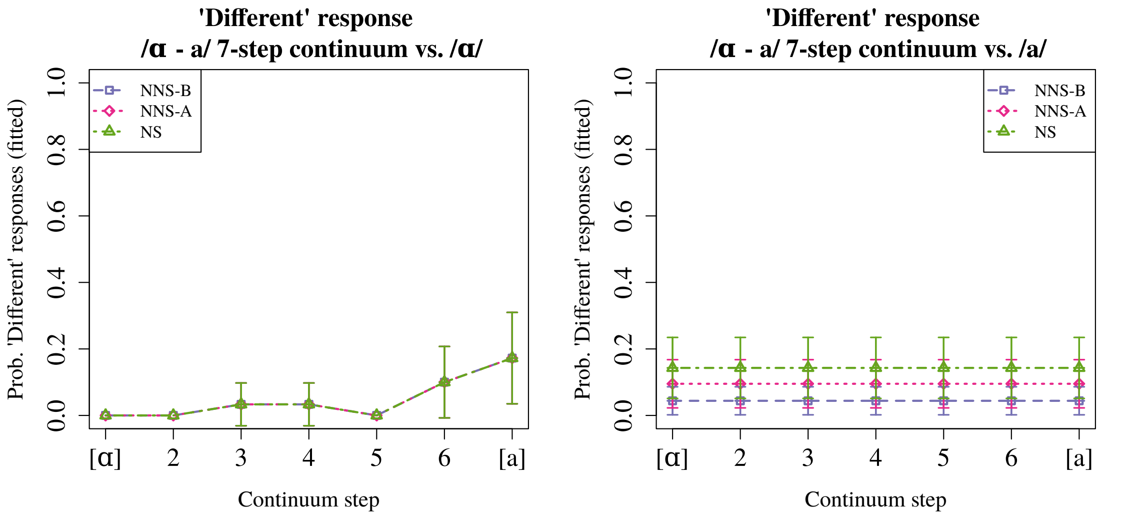


Figure 4.4: Results of AX test, /ɑ - a/

/o/ vowel is acoustically rather far from the token of /ɑ/, and particularly F1-wise, in order to perceptually assimilate: whilst the F1 difference between /ɑ/ and /a/ = 29 Hz, the difference between the former and/o/ = 330 Hz. The F2 difference is slightly smaller in this case, with 252 Hz between /ɑ/ and /a/, vs. 330 Hz between /ɑ/ and /o/.

4.5.2.2 Perception of open-mid back unrounded vowel /ʌ/

The same type of continua as those above were built from the open-mid back vowel /ʌ/: one moving towards /a/ (ED = 182 Hz) and another one moving towards /o/ (ED = 750 Hz). The /ʌ - o/ continuum showed abrupt changes in discrimination, moving from 100% counts of 'same' judgements to 100% 'different' counts along the continuum in all groups and in both directions, which suggests that, just as in the /ɑ - o/ continuum, subjects of all groups attest two different categories to which the formant values of /ʌ/ and /o/ are being mapped (Figure 4.5).

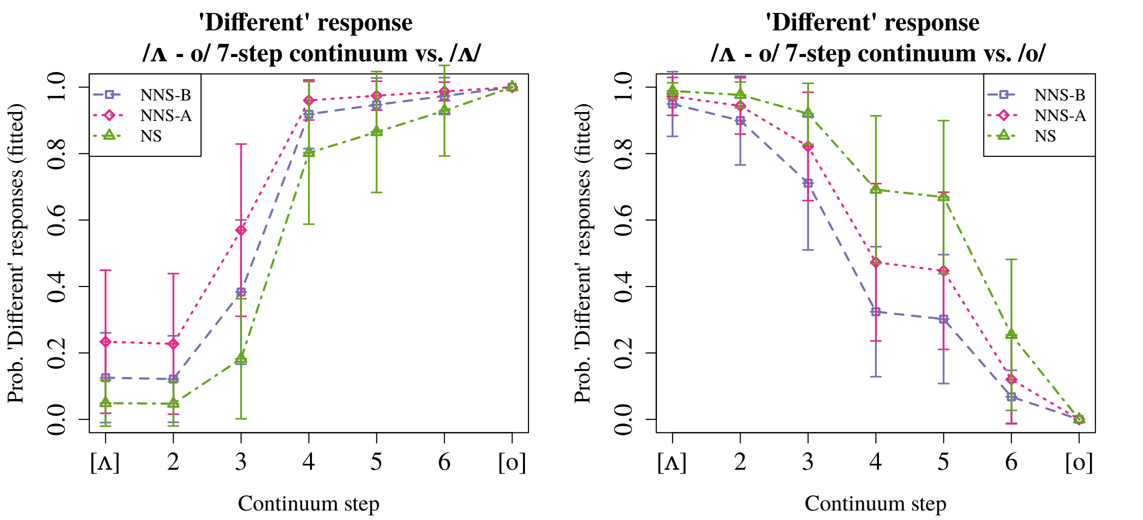


Figure 4.5: Results of AX test, /ʌ - o/

In this case a sharp boundary between the two categories was observed: all groups showed an abrupt change in 'different' counts when compared to /ʌ/, located between stimuli 3 and 5 in the case of native speakers, and between stimuli 2 and 4 in both groups of nonnative speakers. When stimuli were compared against /o/, the discrimination patterns were less abrupt. The models with best fits show that the stimuli pair is a better predictor ($p < 2e\text{-}16$) than the proficiency level ($p = 0.0217$ when discriminated against /ʌ/, and $p = 0.0104$ when discriminated against /o/).

On the other hand, the /ʌ - a/ continuum showed very low discrimination between stimuli in all groups, with 'different' counts that did not reach 40% in both directions. (Figure 4.6). Note that perceptual assimilation of both /ʌ/ and /ɑ/ to the native category /a/ in the nonnative groups is equally strong, regardless of the acoustic distance between these continua endpoints: both ED = 182 Hz and ED = 253 Hz trigger similar levels of low discrimination in all groups. Furthermore, while the F1 difference of 29 between /ɑ/ and /a/ vs. 121 Hz between /ʌ/ and /a/ could explain lower discrimination along the /ɑ -a/ continuum, the F2 of 252 vs. 136 could have predicted otherwise, that is, enhanced discrimination in the /ɑ - a/ continuum and lower discrimination in the /ʌ - a/ continuum. Likewise, an analysis of deviance only shows a very low significance for the stimuli as a predictor ($p = 0.0542$), but the regression does not give any significant values for any stimulus. These results point to the possibility that acoustic distance by itself is not necessarily a decisive factor in discrimination of vowel stimuli.

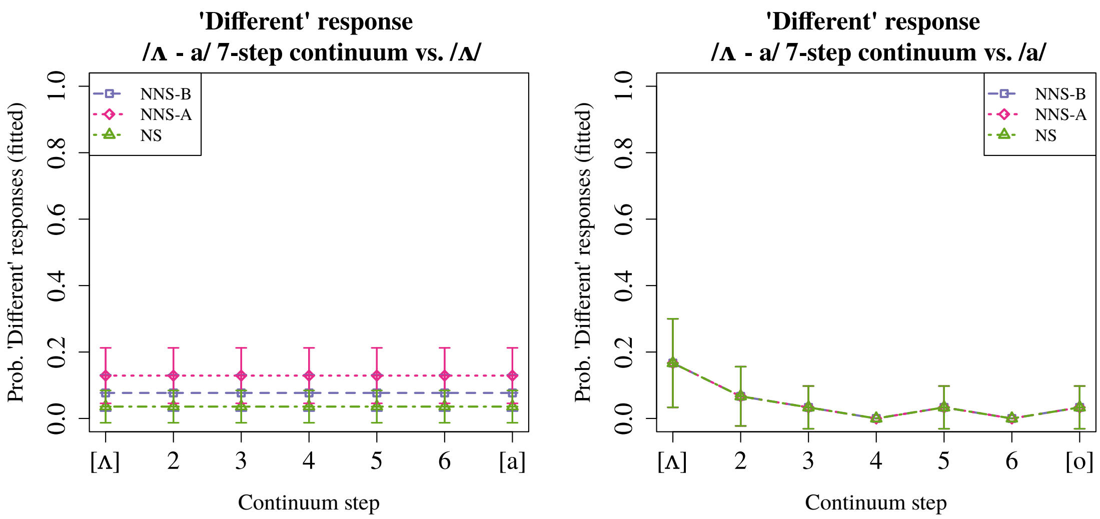


Figure 4.6: Results of AX test, /ʌ - a/

In sum, the results of these four continua combined seem to show two main phenomena. Firstly, they show that both /ɑ/ and /ʌ/ are perceived as acoustically more similar to /a/ than to /o/, regardless of the subjects' proficiency level. This points to the hypothesis that Euclidean distance should be sufficient to explain the warping of non-native onto native values. On the other hand, it seems that regardless of the differences in ED between /ɑ/ and /a/ versus /ʌ/ and /a/, the two

English vowels are perceptually assimilated to Spanish /a/ with similar (if not the same) levels of discrimination along the continua.

4.5.2.3 Perception of the /ɑ - ʌ/ continuum

The results corresponding to this continuum showed a different pattern in discrimination for each group. Native speakers showed a clear boundary effect between tokens 4 and 5 when comparing them against /ɑ/, and at token 4 when compared to /ʌ/. Additionally, and as expected, NS showed a high discrimination rate between endpoints, which was previously shown by the sensitivity measures. Regarding the NNS-A group, token 5 was also the point where the discrimination rate raised when comparing the continuum tokens against /ɑ/, but the group could not reach a total of 100% of ‘different’ responses between endpoints like native speakers did. This incomplete pattern was predicted by the sensitivity measures, which showed to be lower than that of NS. The discrimination pattern shown by beginners was different, with only around 40% of the ‘different’ counts when reaching the point of maximum difference between tokens (i.e. comparison of both endpoints) and zero counts between tokens 1 and 5. However, when the continuum tokens were compared against /ʌ/, beginners and advanced L2 speakers did not show substantial differences in the shape of the discrimination curve, both of which presented smooth patterns and no ceiling discrimination for endpoints, which was again anticipated by the sensitivity measures.

As with the previously analysed continua, the data corresponding to discrimination against /ɑ/ was modelled with a logistic regression in three different ways: one with proficiency level and stimulus as single predictors, a second one with an interaction between level and stimulus, and a third one with both level and stimulus as main but independent predictors. The analysis of deviance showed the highest p-values for a model with level and stimulus as independent predictors (i.e. no interactions). The model used stimulus 4 as a reference level for the stimuli type variable, given that the comparisons against endpoints have extreme values; and NNS-B as the reference level for the proficiency variable. The regression gave significant p-values for stimuli 5 (p=0.007), 6 (p=0.0002), and 7 (p=6.56e-06); likewise, the p-values for NNS-A (p=0.0209) and NS (p=3.47e-06) were also significant. As for discrimination against /ʌ/, the same analysis was carried out, and again, the stimulus and level model as independent predictors (p = 3.487e-15 and p= 2.026e-07, respectively) was the best fit. This model yielded significant results for tokens 1 (p=0.006), 2 (p=0.02), and 6 (p=0.02); however, here NNS-A did not perform significantly different than NNS-B (p=0.08), but the difference between NNS-B and NS was highly significant (p=1.76e-05).[4]

As hypothesised, during the initial state of acquisition the /ɑ/ - /ʌ/ distinction

[4] See Annex 3 for more details.

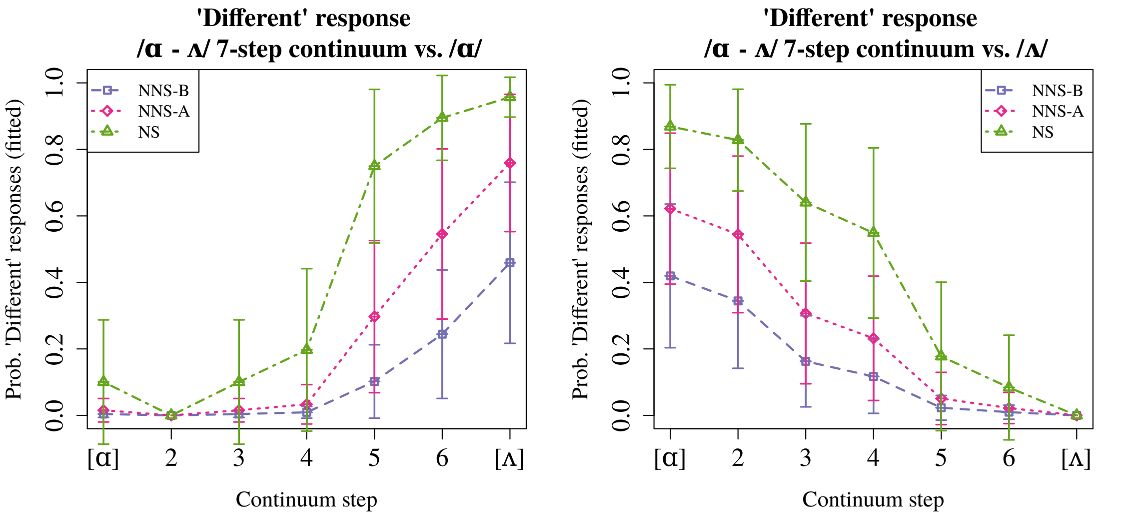


Figure 4.7: AX test, /ɑ - ʌ/: fitted values and confidence intervals at 90%.

is blocked by a perceptual assimilation phenomenon, where these two non-native categories are being perceptually warped to the native Spanish category /a/. But, what is the relation between these two non-native categories in terms of representation? Discrimination along the /ɑ - ʌ/ continuum shows that NNS-B subjects do not reach a level of discrimination above chance at any point along the continuum; at most, they move from a state of certainty in that tokens are the same, to a point of relative uncertainty where they *might* be different. On the other hand, NNS-A subjects reach a point in perception where they know *with some certainty* that tokens are different. However, as the sensitivity measures showed in 4.5.1, and according to the GLM, this level of certainty is still not comparable to that of the NS group; this suggests that even though NNS-A subjects might have left the initial state, they are unable to create a clear categorical distinction between these two L2 sounds.

4.6 General discussion

4.6.1 An initial state: perceptual categories in inexperienced L2 speakers

The results presented above have shown that the predictions in Section 4.3 regarding the initial state have been borne out: endpoint sensitivity measures and discrimination along the continuum suggest that less experienced learners of L2 English with Spanish as L1 perceive both English vowels /ɑ/ and /ʌ/ as members of the L1 category /a/. Figure 4.8 shows an initial state in which L2 speakers categorise English vowels /ɑ/ and /ʌ/ as tokens of the Spanish native category /a/, as predicted. While these two English vowels are acoustically closer to the Spanish vowel /a/ and other initial mappings are unlikely, the chance that one of them could

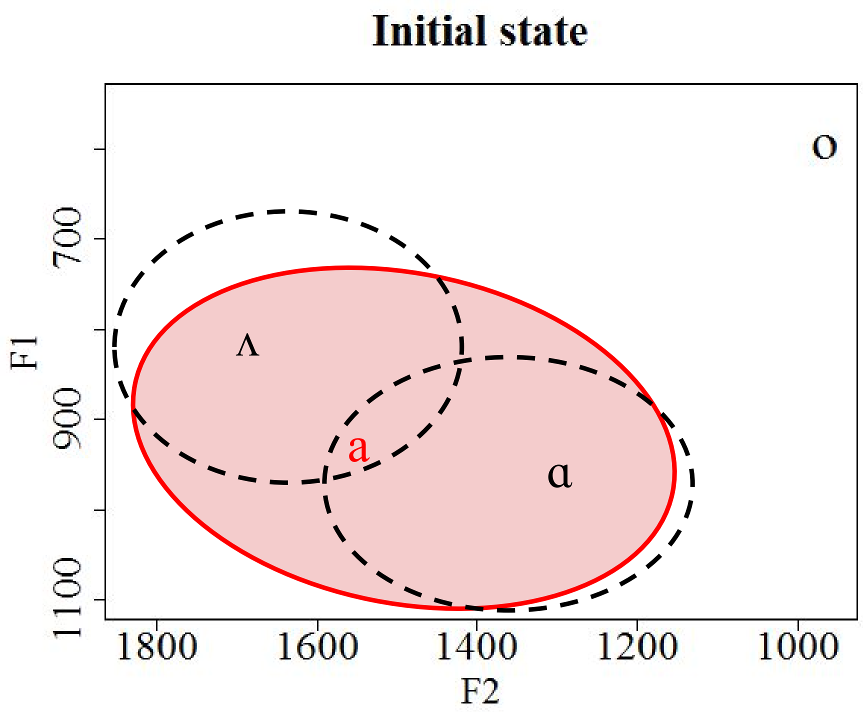


Figure 4.8: Initial state: NNS-B (solid red line) versus the NS group state (black, dashed line).

be initially mapped to another neighbouring L1 category, namely /e/ or /o/ was also possible; however, this was not the case.

Both the endpoint sensitivity measures and the continuum-to-endpoint discrimination show that the level of perceptual assimilation is complete: NNS-B show very low sensitivity between /ɑ/ and /a/, and between /ʌ/ and /a/.

It is clear that even though both English categories share a phonological feature with the Spanish category /o/ (namely, [+back] with /ɑ/, and [-high] [-low] with /ʌ/), neither of them had been perceptually assimilated to it; furthermore, discrimination along /ɑ - o/ and /ʌ - o/ continua shows steep lines between endpoints, with abrupt increases in the 'different' counts between tokens 3 and 5.

However, the /ɑ - ʌ/ continuum shows that despite this 2-to-1 mapping, NNS-B subjects *are* able to perceive a difference between these two vowels, and that they can also perceive a difference as the continuum stimuli move away from the reference vowel. However, their performance barely reaches chance level only at the end of the continuum; before that, their responses show that they are quite certain that these are tokens of the same vowel. Probabilistically speaking, this means that being able to perceive a difference as a non-native speaker with few experience in the L2 is unlikely and may depend on intra-speaker factors that are not measurable (motivation, L2 use, etc.). This includes a wide range of individual differences, some of which are not easy to isolate.

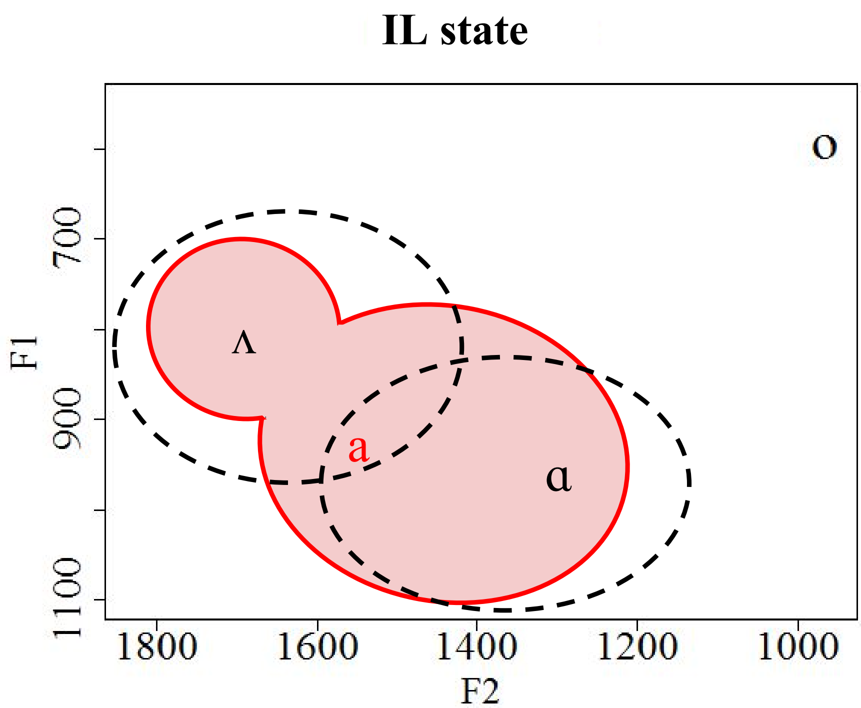


Figure 4.9: An IL state: the acoustic values of both English categories /ʌ/ and /ɑ/ are acceptable values for the same Spanish category /a/, but different distributions allow L2 listeners to discriminate between them.

4.6.2 An IL state: perceptual categories in NNS-A

Likewise, predictions regarding the IL state in NNS-A speakers were also borne out. This non-initial state observed here will be described on the basis of two main points. Firstly, the NNS-A group did not behave in the exact same manner as NNS-B, which proves departure from a non-initial state in L2 phonology acquisition; hence, the initial state shown in Figure 4.8 would not apply; such a state is better described by Figure 4.9. Furthermore, this learning is significant insofar as it reaches an above-chance level, thus becoming qualitatively different from the NNS-B group. However, both groups present great amounts of perceptual assimilation in native-nonnative vowel continua; again, this is expected since the NS group showed similar behaviour.

Secondly, the NNS-A group did not show the same discrimination pattern as the NS group. As mentioned in section 4.4, NNS-A subjects have achieved a rather high proficiency level in the L2; however, this level is given by *performance* in the L2 as measured by standardised tests, whose listening tasks are unable to provide information on a theory of *competence* in perception.

These two findings allow me to state two main claims. The first claim is the rejection of two learning scenarios. Paradoxically, it is not the no-learning scenario: this discrimination task has not allowed us to look into the category system of the IL phonology so as to assert with certainty that phonemically speaking, homophony in minimal pairs such as *cop - cup* no longer takes place. Nevertheless, what the task has allowed us to see is that the IL state of NNS-A speakers is not the same as the initial state, as discrimination increases with linguistic experience[5]. However, it is also clear that a full phonemic split is not the case, as this would have caused the NNS-A group to show the same responses as the NS group. Furthermore, the results have also shown that the NNS-A group has not moved to a state of Two-Category assimilation, or 2-to-2 mapping, in order to gain a contrast between /ɑ/ and /ʌ/. Otherwise, discrimination between either the /ɑ/-/a/ or the /ʌ/ - /a/ vowel pairs would have shown the presence of two different categories: if for instance we were to assume that NNS-A were taking into account the backness of /ɑ/ more than its height so that it could be perceived as /o/ (which was described as a deflection case 1 in section 3.3.2.2), then a flat horizontal line close to the X-axis in the /ɑ - o/ continuum would had been observed.

The second point relates to the possibility that this IL state described above corresponds to the phonetic domain, and does not permeate the phonology layer enough so as to allow for the creation of a new phonemic category. Had the NNS-A group acquired perceptual categories with the same phonemic status as those of native speakers, then both the lines of the discrimination task between tokens of the /ɑ - ʌ/ continuum and its endpoints, and the sensitivity measures between /ɑ/ and /ʌ/ would have been similar. Instead, we found an in-between state, where endpoint sensitivity and discrimination are higher than those of subjects in the NNS-B group, but lower than NS. This probabilistic behaviour is crucial in order to describe the nature of these IL categories, an aspect that we will develop in more detail in the following chapters.

4.6.2.1 The role of experience in noticing

A relevant finding of this experiment is the increase in the ability to discriminate between sounds that are initially perceived as being two instances of the same sound. Increased experience in the L2 has shown to have an effect in a task that attempts to elicit phonological information. In this regard, it seems that more experience triggers further noticing of the acoustic differences between /ɑ - ʌ/; however, this does not mean that (a) any phonemic intake is taking place; and therefore that (b) that this distinction is available in perception of natural speech. At this point it is worth noting that noticing, while strictly required for intake, is not necessarily followed by it. IL categories can then be defined by a movement from no noticing, to noticing the difference between the sounds in question. Nevertheless, this noticing does not seem to be enough in order to (a) create a robust, close-to-deterministic mapping of the acoustic signal onto different parts of the perceptual space; and (b) create a

[5] In fact, the no-learning and subsetting scenarios are not mutually exclusive: homophony can also take place in cases of free variation.

delimited *label* for the areas in the acoustic space that are specifically devoted to receive those mappings. Hence, NNS-A speakers may be able to notice the difference in this task as their attention is specifically focused on perceiving differences, whereas in real-world communication their attentional resources are focused on meaning (VanPatten, 2004). Constant attention to meaning will then result in neglecting the phonetic and phonological components of the input, thus making the learner resort to making use of the robust L1 categories. Despite the situation above, increased L2 input in general seems to somehow ensure processing of the acoustic properties of the L2 phonological inventory, at least to some extent. However, it does not seem to be enough in order to create the new perceptual category, that is, the representational task stated by the L2LP model. While this partial conclusion is being drawn from the absence of the categorical pattern observed in the results of the AX task with tokens of the /ɑ - ʌ/ continuum, a labelling task that confirms the inability to create a category is still needed.

4.6.2.2 Is creation of new categories possible in IL?

Experiment 1 has provided information on the perception of vowel sounds with regard to another vowel sound; that is, whether they are regarded as same or different; this is particularly helpful regarding perception of nonnative sounds and the way in which these are processed by the subjects. However, a discrimination experiment such as this fails to set a straightforward relationship between the outcome of the task and phonological knowledge. What does it mean that the NNS-B group was able to discriminate at chance level between /ɑ/ and /ʌ/? And more importantly, what does it mean that NNS-A subjects were able to discriminate between these two sounds at an above chance level, but without reaching ceiling effects as the NS group did?

The answer seems to be that simply, nonnative speakers are able to perceive the acoustic difference in very specific conditions, and with a higher degree of certainty when the experience with the L2 is higher. Drawing any conclusions regarding the presence of a new phonemic category would be a misinterpretation due to a stretch of the concept of the ability to perceive differences.

However, the fact that more experienced speakers show higher sensitivity to the endpoints of the /ɑ - ʌ/ continuum should not be taken lightly, as it suggests that noticing is enhanced by experience. But again, it does not ensure that intake, i.e. category creation has taken place.

Furthermore, while Experiment 1 is not sufficient to test this hypothesis of partial access to UG in L2 phonology, I would like to set this probabilistic phenomenon observed in the /ɑ/ - /ʌ/ pair as a first argument in this direction. Such behaviour can be explained within the L2LP framework (Escudero & Boersma, 2004) and its division between perceptual and representational task. While it seems to be the case that the NNS-A group has effectively learnt the perceptual task as they are perceiving a certain range of acoustic values in a different manner, the remaining question is: are they remapping the acoustic values to different categories? Are these learners also moving towards the creation of a perceptual category?

In this regard, it is worth thinking of a theory of partial access to UG in which L2 speakers are able to perceive differences in phonological input under certain conditions. This would reject a sensorineural loss hypothesis as a cause for nonnative perception, in the way that Werker and Tees (1984) state. However, what late L2 learners would not have access to, is to create new perceptual phonemic categories after a certain age. Nevertheless, and in order to test this idea, a new experiment would be needed.

4.7 Summary

This Chapter has presented Experiment 1, which attempted to obtain a first impression of changes in the perceptual space in L1 Spanish speakers of English with an advanced level of proficiency, and whether these changes are significant enough to propose that the resulting state is different to that of less experienced learners. The results have shown that even though perceptual assimilation of English vowels to native Spanish vowels is still the same in both groups of non-native speakers of English, there are changes in discrimination of nonnative-to-nonnative vowels, namely /ɑ/ and /ʌ/. However, the level of discrimination in more experienced L2 English learners did not reach that of native speakers, who presented a ceiling effect when comparing against the endpoints of the /ɑ - ʌ/ continuum. This suggests the existence of an IL state, where this raise in discrimination and sensitivity to endpoints does not reach the more robust discrimination patterns of native speakers.

These results, although insufficient, have set the ground for several claims. Firstly, perception of vowels in a Single-Category assimilation scenario seems to work on a probabilistic basis in nonnative speakers. Secondly, input does seem to be important in that some learning takes place, but this learning does not seem to reach the phonemic level. Thirdly, it shows that noticing is enhanced by linguistic experience in the L2, although it is possible that the artificial character of the task could be magnifying an effect that might not be present in situations of naturalistic input. Fourth, the results suggest that even though L2 speakers become increasingly sensitive to perceptual input they could be unable to create new perceptual categories (as described in the subsetting scenario, section 3.3.2.3), for which reason a partial-access hypothesis to UG might be plausible. And finally, the experiment suggests that both deflection and full category split scenarios presented in Chapter 3 should be ruled out.

However, none of the points mentioned above can be firmly posited without further experimentation; in this respect, a labelling task will be necessary, along with a different discrimination task; perhaps more importantly, the relation between these two tasks. Chapter 5 will present Experiment 2, which deals with the questions related to the nature of IL representations.

CHAPTER

FIVE

UNDERSTANDING INTERLANGUAGE REPRESENTATIONS

Perfecto distingo lo negro del blanco
Y en el alto cielo su fondo estrellado

Gracias a la vida
VIOLETA PARRA

5.1 Introduction

The previous chapter has addressed the issue of whether L1 Spanish speakers with more experience in L2 English increase their sensitivity to vowel contrasts that are not present in the L1 and whether they modify their perceptual space through input in such a way that their perception of acoustic differences between /ɑ/ and /ʌ/ becomes more similar to that of NS than to that of less experienced NNS. The experiment performed showed that advanced speakers of L2 English with L1 Spanish have increased their ability to perceive differences along the /ɑ - ʌ/ continuum; however, the NNS-Advanced (NNS-A) sensibility to this vowel contrast is not as robust as that of NS, who displayed an S-curve when stimuli became acoustically further apart in the continuum and a ceiling effect when discriminating prototypical tokens of /ɑ/ and /ʌ/. Furthermore, the results showed that even though tokens of /ɑ/ and /ʌ/ are perceived as members of the L1 /a/ vowel, these two L2 vowels are perceived as different when compared against each other. This suggests that NNS-A have an in-between perceptual category, that is neither the supercategory /a/ that NNS-B seem to have nor the clear categorical distinction made by NS.

This chapter aims to further explore the issue of the nature of these representations, given that the previous discrimination task showed that NNS-A subjects achieved a certain level of robustness that still did not reach that of native speakers. It is worth asking, then, whether the same robustness could be found in other perceptual tasks, especially identification; likewise, the chapter explores this categorisation process in terms of the label used and whether such categorisation departs from what could be considered L1 categorisation.

Here I present Experiment 2, which includes several perceptual tasks that aim to provide information about the presence and nature of perceptual categories in L1 Spanish speakers of English: labelling (both with L1- and L2-like labels), 1-step and 2-step discrimination (AX model), and prototype rating (also in L1 and L2). While Experiment 1 provides important data about the perception of acoustic differences along the /ɑ - ʌ/ continuum and the way in which these vowels are mapped onto the L1 vowel inventory, Experiment 2 focuses on the question of whether the perceptual learning achieved by advanced L1 Spanish learners of English can be considered as creation of a new perceptual category. The results suggest that, despite the evidence towards perceptual learning, the cause cannot be attributed to creation of a new perceptual category; rather, advanced L2 English learners still use their L1 categories, which are nevertheless modified due to increased experience with the L2.

This chapter unfolds as follows: it states the motivation for the experiment, explains the methodology employed, and then presents the results before moving on to the discussion. The discussion addresses the results with a specific emphasis on the nature of the perceptual representations that L2 speakers of English with Spanish as L1 acquire having attained a high proficiency (according to standardised tests). The discussion will take into account three aspects of perception of the /ɑ/ - /ʌ/ contrast. First, it will consider the obtained categorisation data as evidence in favour of a representational task having taken place: that is, whether L2 speakers have created native-like categories to which the incoming input is mapped. Second, it will analyse the relation between discrimination and labelling, with the corresponding implications for a theory of L2 vowel perception. And third, it will take the prototypicality ratings obtained into account as a way to explain the differences between native and nonnative speakers of English in warping of the perceptual space.

5.1.1 Perceptual categories in L2

While Chapter 3 has presented four possible outcomes for perceptual learning, Chapter 4 has shown that L2 learners change their perceptual patterns as a result of increased experience in the L2, thus making it clear that further experience in the TL causes a departure from the initial state. However, the categories that L2 speakers potentially create might not have the same status in the phonology as the native speakers' perceptual categories found in the TL corresponding to the *phonemic* level, and because of which native speakers can distinguish minimal pairs. Hence, and in terms of the *phonological status* of these categories, three options are possible:

- The L2 learner is reusing the same categories of the L1 inventory (as in the no-learning and deflection outcomes described in Chapter 3);

- The L2 learner has created a new phonemic category (as in the full category split outcome);
- The L2 learner has an L1-based phonetic category (as in subsetting), which involves L1 category reuse *and* modification for L2 perception.

Thus, and from now on, a different question remains: does the perceptual pattern observed in Chapter 4 respond to creation of a phonemic perceptual category, or to the fact that a discrimination task allows for storage of medium-term sound representations, which makes acoustic differences more noticeable to nonnative speakers?

Finding such evidence implies looking into a different process: the creation of a new perceptual category, or in Escudero's words, the *representational* task. From an experimental perspective, listeners with two different perceptual categories should be able to identify the tokens from a continuum between vowels A and B as members of two different categories, where the vowel in the middle of the continuum would have a 50% chance of being labelled as either A or B. Importantly, in an identification task the listener needs to do more than just notice a difference: she must have created a long-term abstract category that can be associated with a label and can be readily available whenever incoming input with the specific acoustic values that such category allows is perceived. Hence, an L2 speaker who has acquired new phonemic perceptual categories for the /ɑ/ - /ʌ/ contrast should be able to categorise the first two or three tokens of the continuum between these two vowels as /ɑ/; likewise, she should be able to categorise the last two or three stimuli from the same continuum as /ʌ/.

However, failure to meet this pattern in performance (thus confirming the lack of target-like phonemic categories) cannot provide information on the way in which experienced L2 learners of English deal with the acoustic signal, and how their perception differs from that of less experienced L2 learners. Hence, if what L2 learners do is L1 category reuse and/or modification, it is then worth looking into the manner in which these categories are being reused so as to distinguish their system of representations to that of less experienced L2 speakers of English. Thus, and as mentioned in Chapter 3, we will explore the categories in the L2 learner's IL through several tasks, looking for a consistent account of the nature of perceptual categories in L2 speech perception.

5.2 Linking perceptual tasks with phonological theory

Since the relation between the theoretical aspects of L2 speech perception and the tasks involved in this experiment is not straightforward, I will briefly discuss the specific way in which a certain task contributes to a further understanding of the listeners' learning state and how this learning state allows for a description of the phonological knowledge acquired – which I will ultimately refer to as IL representations. Here I present the expected behaviour of the subjects, the assumptions that sustain these expectations, and the interpretation of such behaviour in terms of phonological knowledge.

5.2.1 L2 labelling: testing for a final state

While Experiment 1 explicitly avoided a labelling task as a manner to observe the presence of any possible non-target representations, Experiment 2 attempts to look specifically for the existence of a target-like representation corresponding to the vowel contrast in question. This is, in fact, a simple way to ensure that L1 Spanish speakers of L2 English have learnt (or not) an abstract representation of the sounds in question, and whether they do it in a consistent manner, comparable to native speakers of English. More consistency in categorisation implies the existence of a robust, readily available representation of a specific vowel that works as a reference to contrast the specific input received while performing the task.

Therefore, it is expected that native speakers, given a certain vowel sound extracted from a continuum between two vowel categories, and two labelling options corresponding to the endpoints of the said continuum, will accurately label the stimuli corresponding to the endpoints as members of the corresponding vowel category. It is also expected that stimuli that correspond to the mid-zone would be categorised randomly (i.e., with around 50% accuracy) as they happen to be on the boundary between the two categories. Such is the behaviour observed in the categorical perception tests performed by Liberman et al. (1957) (see Chapter 2) which nevertheless included consonant stimuli; however, the same effect has been observed in vowels, though to a fewer extent (Pisoni, 1973; Fry et al., 1962)

On the other hand, NNS with little knowledge of English are expected to perform at chance level all along the continuum, as these two categories are not present in the listener's vowel inventory. Conversely, if we present one of these vowels and another English vowel that can be mapped to a different Spanish category (such as /ɛ/, which I predict will be mapped onto Spanish /e/), then it is expected that the listener will display high accuracy at labelling the endpoints, and near-chance behaviour when hearing stimuli that are closer to the boundary between the categories.

Finally, NNS with high proficiency in English are expected to present an in-between behaviour, with better performance in labelling than NNS with little experience in English, but still without the ceiling effect that NS are expected to show.

5.2.2 L1 labelling: understanding the IL perceptual space

In this task, subjects will be asked to use the L1 phonological inventory to categorise the same stimuli as those used in the L2 labelling exercise, i.e. tokens of the /ɑ - ʌ/ continuum. This task will provide information regarding (a) whether NNS show a more consistent pattern in categorisation when using their native categories; (b) whether any of the two vowels in question are considered good exemplars of the L1 categories that are being used for labelling the stimuli, and (c) if there is evidence of perceptual learning in the form of remapping onto other L1 categories. Since this task was not performed by NS of English, the only possible across-group comparison is between less experienced versus proficient learners of English with Spanish as L1. Such comparison is expected to provide evidence regarding noticing: less experienced learners are expected to have a higher count of categorisation into /a/ along the entire continuum versus an around-chance level of categorisation by the end of the continuum in experienced learners.

The L1 labelling results can be interpreted as a way to understand the relation between L1 and L2 categories. While an IL phonology might still retain the labels of the L1, a stronger preference for labelling tokens of the continuum endpoints as different L1 categories (as in the deflection scenario) would show that listeners are noticing a difference that can be interpreted as a change in prototypicality of either one of these endpoints; this can be explained by the fact that experienced listeners are able to perceive a within-category difference, as suggested by Experiment 1. Assuming then that /ʌ/ is a less prototypical exemplar of /a/, it is expected that L2 listeners will categorise /ʌ/ and its surrounding tokens with a lesser level or certainty, thus reaching the aforementioned chance level between /a/ and /o/. In this case, the listener, due to this noticing process, will change the automatised decision-making mechanism given by the L1 and give these boundary-zone tokens the benefit of doubt: a token of this kind can be perceived as an odd member of a certain category, but can also be perceived as an odd member of the neighbouring category. In effect, such behaviour is an indicator of a change in the boundaries of the L1 categories, and a possible intake in the form of a boundary vowel in the L2 that is neither /a/ nor /o/, but instead has no label.

5.2.3 Discrimination: is it bound to the existing perceptual categories?

Discrimination is considered a perceptual task that taps phonetic knowledge as it does not necessarily need to invoke perceptual categories (Colantoni et al., 2015); here, categories are expected to perform either an indirect role, or none. Despite what earlier Categorical Perception experiments have shown regarding how discrimination depends on phonemic categories (Liberman et al., 1957), it has also been shown that such behaviour is not common for vowels and that whenever CP takes place, it is ultimately an effect of the task (Gerrits & Schouten, 2004). While I am not taking a theoretical stance regarding this matter, the goal with this task is to compare NS discrimination ability against that of both NNS groups. I predict that NS will show a peak in discrimination at the boundary (however mild) between the two categories, and that NNS would not show such pattern. In fact, it is expected that NNS will show very low levels of discrimination, with a *d'* close to zero. What NNS with high proficiency in English should present is higher sensitivity with no peaks, which falls in line with higher ability to notice differences, but without the peak that is assumed to show due to the presence of a category boundary.

5.2.4 Ratings: Prototype theory in perception

According to prototype theory (Rosch, 1999), categories have fuzzy boundaries, and they are better conceptualised in a radial manner (that is, with elements that are closer to the prototype and thus are considered to be better-fitting exemplars than the peripheral elements which are considered dubious exemplars of the category) than in a logic manner. This approach is particularly suitable for perception – and vowels especially – given the great amount of dispersion of their acoustic values in a multidimensional acoustic space.

However, traditional aproaches in phonology work on a binary basis, and categories are understood as bundles of features that have either a positive or negative value. Hence, a vowel like /ʌ/ can be described in terms of a set of features: [-high], [+back], [-tense], etc. From this perspective, categories should be clearly delimited, and categorical perception effects (Liberman et al., 1957) show that perceptual phonemic categories are not an exception: the perceptual space is modelled by the phonemic categories in such a way that the perception of the acoustic values associated with these features is also influenced by the binarity of the feature system[1].

Furthermore, the Native Language Magnet effect (Kuhl & Iverson, 1995) shows that prototype values of a vowel category 'attract' tokens that are acoustically different but relatively close in the perceptual space, so that they are perceived as instances of the same category. Non-prototype tokens of a category do not create this effect. The Native Language Magnet Kuhl and Iverson (1995) adheres to prototype theory as it looks for the effects that prototypes of speech categories create in perception of other tokens of the same category. Thus, in an AX discrimination task between a stimulus considered by listeners to be a prototype of a perceptual category and another stimulus that is less prototypical but also part of the category, the stimulus that is acoustically closer to the prototype would be perceived as the

[1]While the real acoustic values are rather scattered in the acoustic space with the subsequent overlapping of values between categories, the resulting fuzzy boundaries do not imply that there are no boundaries.

same sound and, therefore, both stimuli will be considered members of the same category (i.e. a discrimination task between a prototype and a perceptually close token will get a 'same' judgement by a listener). Conversely, poor exemplars of the category do not 'attract' the surrounding sounds in the continuum, for which peripheral tokens of the category are easier to perceive as 'different' in an AX task.

This theory's implication for L2 speech perception is that nonnative sounds that are similar to a native category will be particularly difficult to discriminate as different. This is supposed to be due to being attracted by the prototype of the native category. On the other hand, tokens of nonnative sounds that are considered to be less prototypical regarding the L1 category would be perceived as being different.

The goal of using prototypicality ratings in this experiment is to know whether prototypes of both native and nonnative categories play a role in perception of nonnative vowels that are originally mapped onto a native category. If, as mentioned previously, /ʌ/ is a poorer member of the L1 category /a/ than /ɑ/, discrimination between /ʌ/ and its surrounding tokens should be more successful than discrimination between /ɑ/ and its acoustically closest tokens.

5.3 Hypotheses

1. Regarding the nature of IL representations, I hypothesise that experienced L1 Spanish learners of English **do not** create *phonemic* categories in the L2 in cases of Single-Category assimilation, particularly in the case of the /ʌ/ - /ɑ/ contrast found in English; rather, they become able to perceive differences between these two vowels while still being identified as exemplars of the same L1 category /a/. This can be observed in the following:

    - An identification task along a /ɑ - ʌ/ continuum with L2-like labels showing a random categorisation pattern, both in experienced and less experienced L2 learners: that is, that nonnative speakers of whatever level of experience in the L2 will be equally likely to choose either the /ɑ/ or /ʌ/ label. Conversely, native speakers of English will show a change in their identification preferences, from /ɑ/ for the tokens that are closer to the /ɑ/ end of the continuum, to /ʌ/ for those that are closer to the /ʌ/ side of it.

    - A discrimination task between adjacent tokens of the continuum that shows low discrimination, with no peaks, when comparing with tokens that are across the phonemic boundary set by native speakers. I predict that discrimination should be higher overall in experienced L2 speakers than in those less experienced; and that native speakers of English should have low discrimination between tokens that were categorised as members of the same category in the labelling task, and higher when comparing tokens across the phonemic boundary.

- An identification task along a /ɑ - ʌ/ continuum with L1-like labels that categorises all the tokens in the continuum as one of the L1 categories. I predict that nonnative speakers of all levels will choose /a/ (instead of /o/) all along the continuum.
- A rating task for each token along the /ɑ - ʌ/ continuum, applied after subjects have categorised them into L2-like labels, where nonnative subjects give poor ratings to all tokens. Native speakers are expected to give higher ratings to the endpoints and lower ratings to boundary stimuli.
- A rating task for each token along the /ɑ - ʌ/ continuum, applied after subjects have categorised them into L1-like labels, where more experienced L2 subjects give poor ratings to all tokens. Native speakers are expected to give higher ratings to the endpoints and lower ratings to boundary stimuli.

5.4 Methodology

5.4.1 Task 1: L2 labelling and rating

5.4.1.1 Subjects

Subjects were divided into three groups: a control group with native speakers of American English (NS, N=10), and two experimental groups of native speakers of Chilean Spanish, with beginner (NNS-B, N=11) and advanced (NNS-A, N=9) knowledge of English. Advanced speakers had at least one immersion experience in an English-speaking country after they were 18 years old or were abroad for a shorter amount of time, but had a C1+ score in a standardised proficiency test (TOEFL or IELTS). Subjects in the NNS-B group had no previous immersion experiences and either had below B2-level scores in proficiency tests or no test score and a low self-assessment score.

5.4.1.2 Stimuli

Formant values for the stimuli were taken from recordings made by a native speaker of American English reading a list of words that contained the English phonemes /ɑ/, /ʌ/, and /ɛ/ in a C_C context and the carrier sentence *I say ____ once*. For stimulus recording, the American English and the Chilean Spanish speaker were recorded in a sound-attenuated booth using an AKG condenser microphone, model C520. The sounds were recorded on a HP ProBook 6570b with the Praat recording interface. The Spanish speaker was recorded in a quiet room, using a handheld microphone on an iMac computer with the Praat recording interface. Both recordings were made at a sampling frequency of 44100 Hz. The vowels were then extracted from the rest of the words and rated by 10 native speakers of American English, on a scale from 1 (very poor exemplar of the category) to 5 (very good exemplar of the category). Formants of the tokens that obtained the best ratings were measured and averaged. These values were taken to generate a 5-step continuum of synthesised vowels, which were created using the Praat Klatt synthesizer (Boersma & Weenink, 2012). Vowels /ɑ/ and /ʌ/ were the endpoints of the first continuum, and vowels /ʌ/ and /ɛ/ the endpoints of the second continuum.[2].

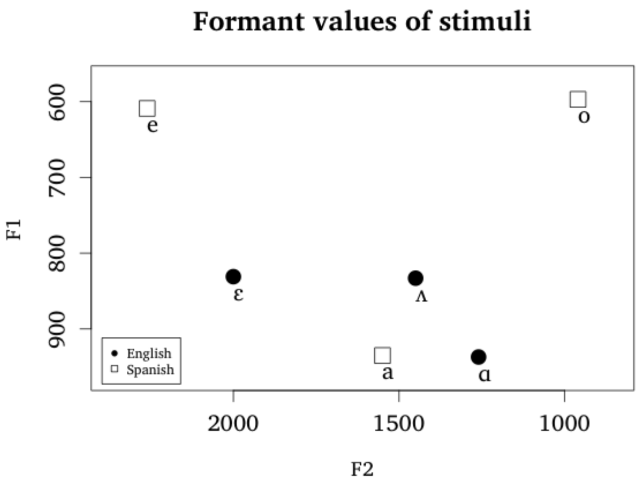


Figure 5.1: Vowel stimuli (black) and Spanish reference vowels (white).

5.4.1.3 Procedure

The experiment was carried out online using the application Limesurvey. Subjects were recruited online through social networks and were given a unique access key to the experiment. After logging in, they gave their informed consent and responded to a short language background questionnaire (see Appendix 2). Unfinished responses were discarded.

For this task, subjects were asked to label the auditory stimuli as English vowels. In order to minimise orthographic biases, pictures that referred to words containing a specific vowel were used as labels. First, subjects were exposed to a short explanation block with three pictures and their corresponding audio file with the word being uttered by a native speaker of English: *bed*, *pot* and *cup*. Each question presented one stimulus at a time, which corresponded to one of the resulting tokens of the continua created. The subjects listened to the stimulus and were then asked to choose between two pictures, with the instruction: “Listen to the vowel. The vowel that you just heard is like the one in the word...”. The alternatives given corresponded

[2]I predict that the /ɛ/ vowel will be mapped onto the Spanish category /e/, for which all three groups of speakers should be able to map /ʌ/ and /ɛ/ tokens onto different perceptual categories.

to the endpoints of the continuum to which the token belonged, so that a token from the /ɑ-ʌ/ continuum could be classified either as the vowel in the word *pot* or as that in the word *cup*.

Immediately after subjects chose a picture, they were asked to rate the vowel they heard on a scale from 1 to 5, where 1 = poor exemplar of the vowel, and 5 = excellent exemplar.

All stimuli were presented randomly and one at a time. Each stimulus was repeated 5 times, thus making a total of 50 stimuli.

In order to trigger L2-like perception, all instructions were in English, so that subjects were in the correct language mode (Grosjean, 2001).

5.4.2 Task 2: 1-step and 2-step discrimination

5.4.2.1 Subjects

Subjects were the same as those of Task 1.

5.4.2.2 Stimuli

Stimuli for the discrimination task were the same as those of Experiment 1, but were presented in pairs. One set of pairs consisted of two immediately adjacent stimuli in the continuum (1-step discrimination: 1-2, 2-3, 3-4, and 4-5), which were presented to the listener with an inter stimulus interval (ISI) of 1 sec. The other set of pairs consisted of stimuli that were acoustically 2 steps apart from each other in the continuum (1-3, 2-4, and 3-5), also with an ISI of 1 sec.

5.4.2.3 Procedure

In this task, subjects discriminated between two vowel sounds using an AX paradigm (same - different). Subjects heard a stimuli pair, after which they were asked whether the stimuli heard were the same or different. Stimuli were randomised and repeated five times when both vowel sounds were different, and two times when the pair of sounds consisted of exact identical stimuli. For each discrimination trial where the stimuli were different (e.g. step1 - step 2), The subject also heard trials where the stimuli were the same (step 1 - step 1, and step 2 - step 2, two times each). Instructions were also given in English.

5.4.3 Task 3: L1 labelling and rating

5.4.3.1 Subjects

As above.

5.4.3.2 Stimuli

Same as those of Task 1.

5.4.3.3 Procedure

The second labelling task employs the same stimuli, but the categorisation choices are represented by pictures that refer to Spanish words: *pan* ('bread'), *red* ('net') and *ron* ('rum'). In order to make this task as similar as possible to L2 labelling, pictures of the words listed are used instead of the orthographic form[3]. A short block was provided first, where a recording of the word was provided along with the corresponding picture. Native speakers of English who took the test did not carry out this labelling task. A category rating task was presented to the subjects after they categorised each stimulus: subjects were asked to rate the degree of prototypicality for each one of the tokens on a scale of 1 to 5, where 1 is very poor and 5 is excellent. Stimuli were randomised and repeated 5 times. Instructions for this task were given in Spanish.

5.5 Results

5.5.1 Task 1: L2 labelling

Overall, the L2 labelling results show that both groups of nonnative speakers categorise the English vowels /ɛ/, /ʌ/, and /ɑ/ in a similar manner: while /ɛ/ and /ʌ/ are identified as members of two different categories, tokens from the /ɑ - ʌ/ continuum are mostly identified as members of the same category with a stronger preference for the label with /ʌ/ (i.e *cup*). Table 5.1 shows the obtained categorisation rates for the stimuli corresponding to the /ʌ - ɛ/ continuum; and Table 5.2 shows the proportions in categorisation along tokens of the /ɑ - ʌ/ continuum.

Figure 5.2 shows the fitted values for both continua and the three subject groups, with a 95% confidence interval; standard errors for each continuum step and group can be found in Appendix 3. Considering the binary nature of the responses (2-alternative categorisation), a logistic regression was used in order to model the responses. An analysis of deviance showed that the best fit for the responses along both continua was given by a model that considered an interaction between group (NNS-B, NNS-A, and NS) and stimulus (1 to 5).

[3]While Spanish has a transparent orthographic system for vowels, the pictures were used in order to create consistency with the previous labelling task.

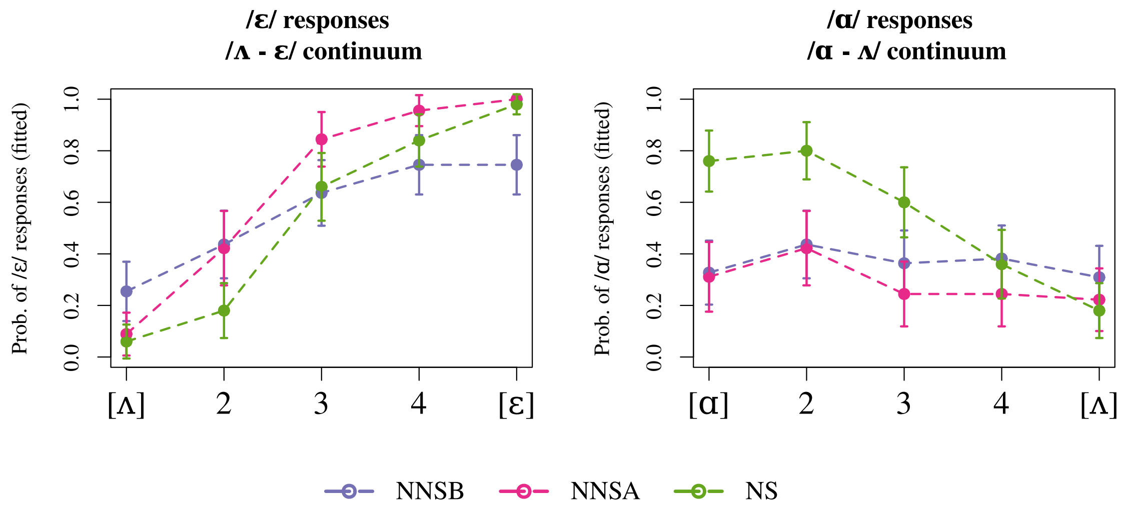


Figure 5.2: Results of L2 labelling task.

5.5.1.1 Less experienced L1 Spanish speakers of English (NNS-B)

The /ʌ-ɛ/ continuum shows a high proportion of counts (74.5%) where English /ʌ/ (stimulus 1 in this continuum) is categorised as the vowel /ʌ/ (i.e. selection of the *cup* picture). The proportion of these counts decrease as the stimuli become acoustically closer to the other endpoint /ɛ/ (stimulus 5 in the continuum), which was mostly categorised as /ɛ/ (i.e. selection of the *bed* picture) with a categorisation rate of 74.5%. While the vowel /ɛ/ is not exactly the same as /e/, subjects had no further difficulty in categorising it as such[4]. This categorisation pattern with a boundary between stimuli 2 and 3 of the /ʌ - ɛ/ continuum shows that the NNS-B group mapped /ʌ/ and /ɛ/ onto two different perceptual categories, namely, /a/ and /e/ respectively; however, their performance in the task suggests more difficulty in establishing a clear boundary between the two categories. The logistic regression confirms that NNS-B responses to the stimuli corresponding to the /ʌ - ɛ/ continuum are significantly different to those of NNS-A; the latter group was much closer to the NS response pattern. Figure 5.2 (left, purple line) shows how the /ɛ/ responses given by NNS-B follow a similar pattern to that of the other groups.

Table 5.1: Categorisation proportions, /ʌ - ɛ/ continuum

Stim.	**1 (= ʌ)**		**2**		**3**		**4**		**5 (= ɛ)**	
Cat.	ʌ	ɛ	ʌ	ɛ	ʌ	ɛ	ʌ	ɛ	ʌ	ɛ
NNS-B	0.745	0.255	0.564	0.436	0.364	0.636	0.255	0.745	0.255	0.745
NNS-A	0.9	0.1	0.55	0.45	0.15	0.85	0.05	0.95	0	1
NS	0.94	0.06	0.82	0.18	0.34	0.66	0.16	0.84	0.02	0.98

Regarding the /ɑ - ʌ/ continuum, the NNS-B group presented very little variation

[4]Traditionally, [ɛ] has been considered an allophone of /e/ in Spanish (Bradley, 2014); specifications for /e/ are [+front], [-low], [-high], [-round].

in categorisations of the tokens along the continuum, which were mostly labelled as instances of /ʌ/ with a proportion of counts that remained relatively stable along the continuum, with values between 56.4% (stimulus 2) and 69.1% (Stimulus 5). The preference for the /ʌ/ label over /ɑ/ suggests that the vowel in *cup* is less ambiguous to map to /a/ than the vowel in *pot*, perhaps due to an orthographic effect that persisted regardless of the experimental design. Unlike the previous continuum, this suggests that subjects are unable to perceive differences along the continuum, which is observable in the lack of the switching pattern: here, the categorisation behaviour remains mostly constant. The fitted values for /ɑ/ responses are plotted in Figure 5.2 (right, purple line) and are shown as a line that runs almost paralell to the X-axis.

Table 5.2: Categorisation proportions, /ɑ - ʌ/ continuum

Stim	**1 (= ɑ)**		**2**		**3**		**4**		**5 (= ʌ)**	
Categ	ɑ	ʌ	ɑ	ʌ	ɑ	ʌ	ɑ	ʌ	ɑ	ʌ
NNS-B	0.327	0.673	0.436	0.564	0.364	0.636	0.382	0.618	0.309	0.691
NNS-A	0.35	0.65	0.475	0.525	0.275	0.725	0.275	0.725	0.25	0.75
NS	0.76	0.24	0.8	0.2	0.6	0.4	0.36	0.64	0.18	0.82

5.5.1.2 Experienced L1 Spanish speakers of English (NNS-A)

The NNS-A group was also able to make a distinction between /ɛ/ and /ʌ/, although much more clearly defined than that drawn by the NNS-B group. Here, correct categorisations at endpoints reached over 90%, with a change in categorisations immediately after stimuli 2. This behaviour is strikingly close to ideal categorical perception[5], where a clear boundary is defined by an abrupt decrease of counts at some point along the continuum and increase of counts of the label corresponding to the other end of the continuum. The fitted model shows that NNS-A behaviour in this continuum is significantly different from that of NNS-B, but not from NS. Figure 5.2 (left, pink line) shows the fitted values for /ɛ/ responses.

The /ɑ - ʌ/ continuum, according to the predictions, shows again a clear merging of these two categories, with counts of /ʌ/ that reach around 80%. While there is some fluctuation along the continuum, both categories /ɑ/ and /ʌ/ are being triggered in similar proportions all along the continuum[6]. This suggests that for subjects in the NNS-A group (a) there is no noticing of differences along the continuum, thus confirming that this group of speakers has a similar perceptual pattern to that of the NNS-B group; and (b) that the preferred L2 label for this category

[5]Labelling-wise.

[6]The reason for this preference could be found in an orthographic effect, where both NNS groups (but more definitely NNS-A, who certainly know the spelling of *pot*) associate the ‘o’ in *pot* with a rounded sound, despite having listened to the initial audio file with a clear /ɑ/ and therefore assuming that the vowel must be also rounded in speech.

is, again as in the NNS-B group, *cup*, and not *pot*[7]. The fitted values show that the categorisation pattern corresponding to the NSS-A group (right, pink line) is significantly different form that of NS; on the other hand, the NNS-A pattern is not significantly different from that of NNS-B.

5.5.1.3 Native speakers of English (NS)

Finally, native speakers of English presented a rather categorical pattern in labelling, as expected. Here, the sudden changes in categorisation are present in both continua. The steepest slope in the /ʌ - ɛ/ continuum is also between stimuli 2 and 3 (as in both NNS groups), which suggests that category boundaries are in the same place in the acoustic space, both in native speakers of English and Spanish. Figure 5.2 shows the fitted values for /ɛ/ responses along the continuum (left, green line).

In /ɑ -ʌ/, the NS group shows again a categorical pattern, which is not the case in neither of the NNS groups. Here the line showing the probability of /ɑ/ responses (right, green line) suddenly changes again, this time between stimuli 3 and 4 of the continuum. The absence of ceiling effects at the endpoints of this continuum suggests that the stimuli used were not exactly prototypical for this particular group of NS, which is highly likely given the variety of vowel systems within General American English, and the synthetic nature of the stimuli.

5.5.2 Task 1: L2 ratings

Since ratings are paired with categorisation in a forced-choice task, their informativeness is subordinated to these previous categorisations. In the case of L2 ratings, categorisation results suggested that every token along the continuum was perceived in a similar way, with a preference for the *cup* label over *pot*. Table 5.3 shows the mean ratings assigned to each stimulus in the /ɑ - ʌ/ continuum according to the way in which they were categorised; Table 5.4 shows the mean ratings for the /ʌ - ɛ/ continuum.

Table 5.3: Mean rating scores for each stimulus according to their categorisation, /ɑ - ʌ/ continuum.

Stim	**1 (=ɑ)**		**2**		**3**		**4**		**5 (=ʌ)**	
Categ	ɑ	ʌ	ɑ	ʌ	ɑ	ʌ	ɑ	ʌ	ɑ	ʌ
NNS-B	1.94	2.51	1.75	3.10	1.90	2.66	1.86	2.21	1.41	2.63
NNS-A	2.57	2.77	2.84	2.38	2.64	3.00	2.55	3.12	2.70	2.69
NS	3.24	2.17	3.33	1.90	3.07	2.70	2.83	2.41	2.78	3.02

The initial uncertainty in categorisation is somewhat redundant with regard to ratings; in this respect, a rating task could be more informative in evaluating the

[7] A further reason for choosing /ʌ/ over /ɑ/ may be found in the production patterns of Chilean Spanish: according to Sadowsky (2012), vowels in this system have significantly more centralised values than other varieties of Spanish.

stimuli tokens against their own L1 categories. However, ratings given to stimuli when evaluated as members of L2 categories show how well the given tokens adjusted to whatever perceptual category they had in mind while performing the categorisation task. Additionally, it can be hypothesised that uncertainty in categorisation could reflect on low ratings.

Table 5.4: Mean rating scores for each stimulus according to previous categorisation; /ʌ - ɛ/ continuum.

Stim	**1 (=ʌ)**		**2**		**3**		**4**		**5 (=ɛ)**	
Categ	ʌ	ɛ	ʌ	ɛ	ʌ	ɛ	ʌ	ɛ	ʌ	ɛ
NNS-B	2.15	2.29	2.26	2.21	1.90	2.60	1.86	2.85	1.71	2.93
NNS-A	3.07	2.25	3.12	2.68	1.86	3.13	2.00	3.44	0.00	3.64
NS	3.17	2.33	3.17	2.44	2.82	3.06	1.50	3.40	1.00	3.73

5.5.2.1 NNS-B group

The NNS-B group assigned low ratings to the stimuli along the /ɑ - ʌ/ continuum, which were significantly lower for stimuli categorised as /ɑ/ than for those categorised as /ʌ/. An unpaired, two-tailed t-test[8] yields the values $t(230.93)$ = -5.81, p < 0.001. This pattern suggests that whatever category is being triggered by these continuum stimuli, they seemed to be better fits for the vowel in *cup* and whenever the *pot* label was chosen, then stimuli would be considered worse exemplars of it.

On the other hand, while the NNS-B subjects still showed the categorical pattern in identification along the /ʌ-ɛ/ continuum, they also gave low ratings to these stimuli. However, a t-test shows that, again, one category obtained significantly better scores than the other; in this case, tokens that were categorised as /ɛ/ had higher scores than those labelled as /ʌ/ ($t(268.38)$ = 3.99, p < 0.001). This can be easily noted when comparing the endpoints of the continuum: stimulus 1 in the continuum (=/ʌ/) obtained a 75% of categorisations as /ʌ/, its mean rating was only 2.15; however, stimulus 5 (=/ɛ/), while also obtaining 75% of categorisations as /ɛ/, reached a mean rating of 2.93. This difference suggests that *cup* is still seen as a poor fit for /ʌ/ tokens, which I predict, are being mapped onto the L1 category /a/.

5.5.2.2 NNS-A subjects

Unlike NNS-B, the NNS-A group has assigned higher ratings to the tokens of the /ɑ - ʌ/ continuum and, furthermore, the ratings assigned the categories did not differ significantly ($t(122.75)$= -0.88, p = 0.38) in any case. Hence, while NNS-A categorised these tokens mostly as exemplars of /ʌ/, whenever they assigned a

[8]Unless specified, all t-tests in this chapter correspond to this type.

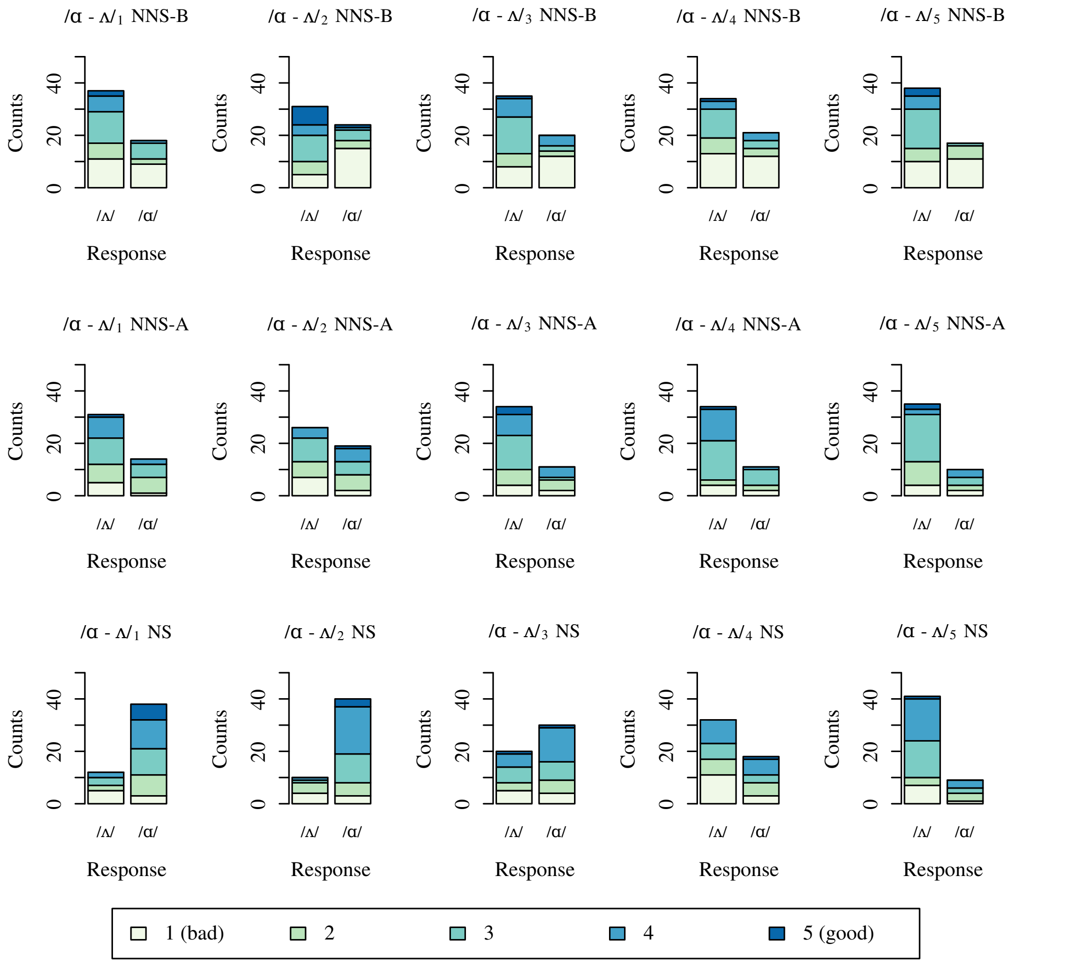


Figure 5.3: L2 rating task results by token, /ɑ - ʌ/ continuum (/ɑ - ʌ/$_1$ = first continuum token, i.e. /ɑ/).

token to the category /ɑ/, the tokens were still rated as good for this category. This suggests that while the categorisation pattern is similar to that of the NNS-B group, they seem to assume that tokens fit both categories equally well.

Tokens of the /ʌ - ɛ/ continuum were also rated as mostlygood members of the corresponding categories by this group, particularly those tokens that were categorised as /ɛ/: while stimulus 5 (=/ɛ/) obtained 100% of the categorisations as /ɛ/ and a mean rating of 3.64, stimulus 1 (=/ɑ/) obtained 91% of categorisations and a mean rating of 3.07. The t-test for significant differences between the means obtained by each category ($t(122.95) = 2.37$, $p < 0.05$) supports this claim.

5.5.2.3 NS subjects

Finally, the NS group rated their tokens roughly according to the categorisation given: while the first stimulus in the /ɑ - ʌ/ continuum (categorised mostly as /ɑ/) obtained a mean rating of 3.2, the last stimulus (categorised mostly as /ʌ/) obtained a mean rating of 3.0. The other stimuli were assigned a lower value. However, there is still an effect in ratings according to the category: a t-test shows that, contrary to both NNS groups, ratings are higher for /ɑ/ than /ʌ/ ($t(235.849) = -3.54$, $p < 0.001$). Such asymmetry can be explained by the rather high F2 value for /ʌ/, which could have caused NS subjects to judge the vowel as non-prototypical.

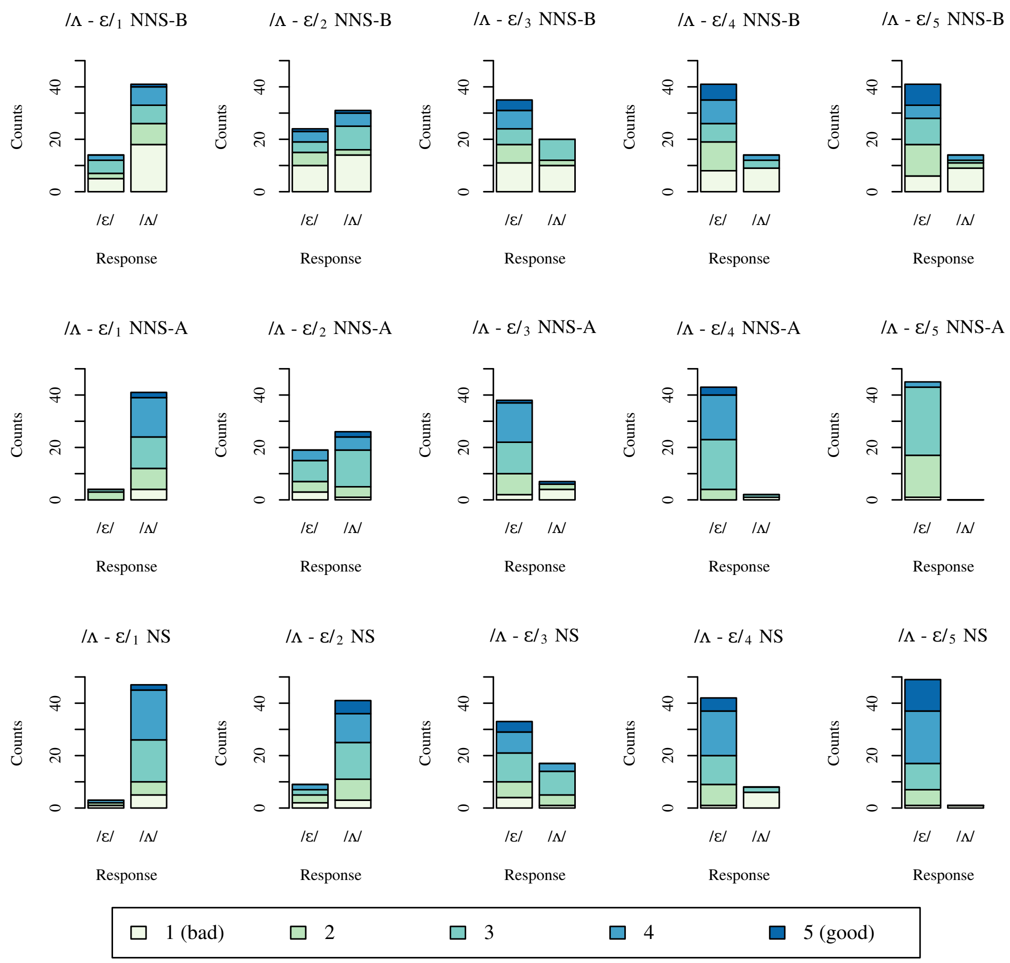


Figure 5.4: L2 rating task results by token, /ʌ - ε/ continuum (/ʌ - ε/$_1$ = first continuum token, i.e. /ʌ/).

5.5.3 Task 2: Discrimination

The *d'* scores for discrimination presented an unexpected pattern: considering the evidence obtained so far according to which Spanish speakers of English map English /ε/ onto /e/ and /ʌ/ onto /a/, the prediction was that sensitivity for the /ε - ʌ/ continuum would be similar among the three groups, and different among groups for the /ɑ - ʌ/ continuum. However, *d'* scores for these two continua were exactly the opposite: discrimination patterns among the three groups were much more similar along the /ɑ - ʌ/ continuum instead. Figure 5.5 shows the mean sensitivity scores obtained by all groups with a 95% confidence interval.

Considering that the amount of repetitions per subject was higher in this discrimination experiment (5 signal trials and 4 noise trials), the *d'* scores were calculated individually, thus making a statistical analysis possible. However, since the sample size was still small and the distribution was not normal, an ANOVA test was still not possible; instead, a non-parametric test (Kruskal-Wallis test) was carried out. Standard errors of d-prime and the results of the Kruskal-Wallis test can be found in Appendix 3.

5.5.3.1 1-step discrimination

The first column in Figure 5.5 shows the *d'* scores of the 1-step discrimination task, plotted against the stimuli pairs.

Sensitivity to adjacent stimuli in the 1-step discrimination task between adjacent tokens of the /ɑ - ʌ/ continuum did not reach high levels in any of the groups, with *d'* scores < 2.00. While the results show different trends in terms of the slope (with the NS group showing the expected peak at the boundary set by the L2-label identification test), no stimuli pair yielded significantly different d-prime scores. Table 5.5 shows the mean d-prime values obtained by each group for each stimuli pair along the /ɑ - ʌ/ continuum.

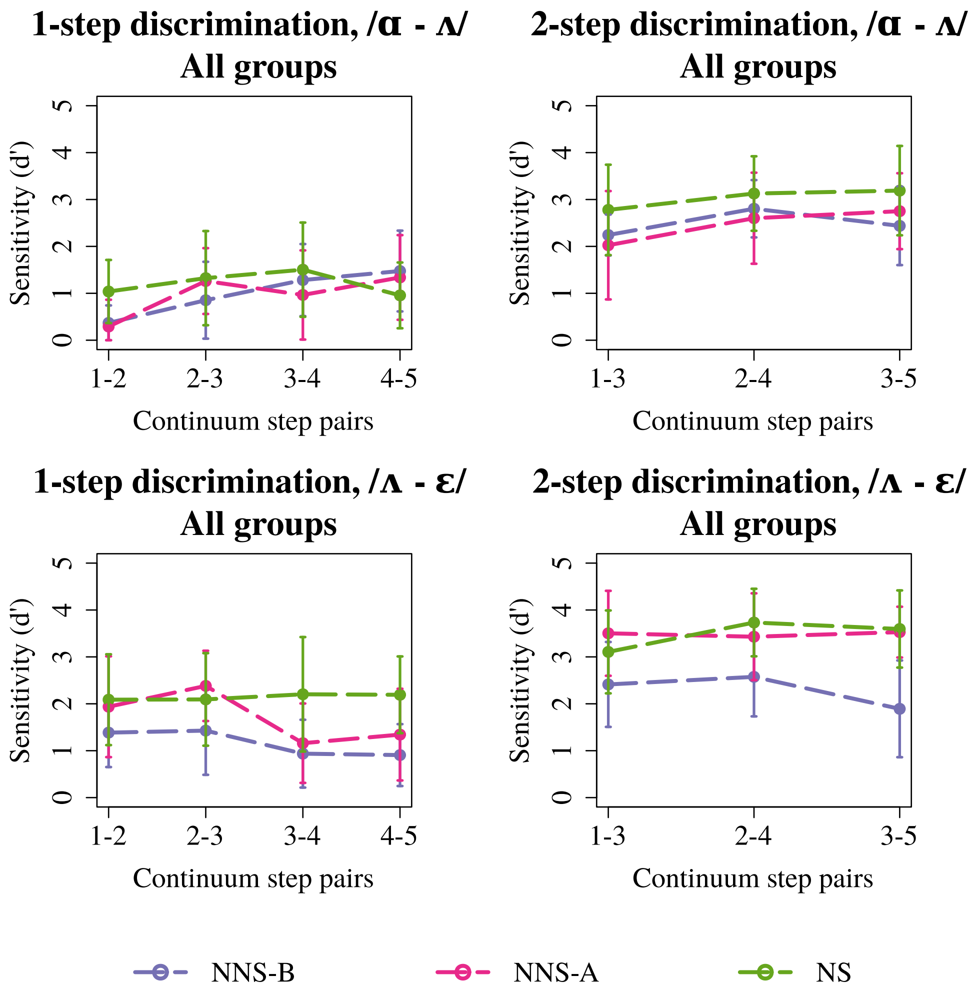


Figure 5.5: Sensitivity (d') scores, discrimination task.

Table 5.5: /ɑ - ʌ/ continuum sensitivity (*d'*) scores, 1-step discrimination

	1-2	2-3	3-4	4-5
NNS-B	0.37	0.85	1.28	1.48
NNS-A	0.29	1.26	0.97	1.34
NS	1.03	1.32	1.34	1.00

The /ʌ - ɛ/ continuum triggered different discrimination behaviour (Table 5.6). The NS group showed overall higher sensitivity than the nonnative groups, but without the peaks predicted under a categorical perception approach: for these subjects, all stimuli pairs were perceived as equally different. NNS-B showed a similar pattern, but with lower *d'* scores. The only group that showed a peak was NNS-A, when discriminating tokens 2 and 3 *d'*= 2.38). Again, the Kruskal-Wallis test showed no significant differences for any of the stimuli pairs.

Table 5.6: /ʌ - ɛ/ continuum sensitivity (*d'*) scores, 1-step discrimination

	1-2	2-3	3-4	4-5
NNS-B	1.39	1.43	0.94	0.91
NNS-A	1.94	2.38	1.16	1.34
NS	2.09	2.09	2.20	2.19

5.5.3.2 2-step discrimination

The 2-step discrimination task yielded higher *d'* scores and slightly different curve patterns for the /ɑ - ʌ/ continuum, and a noticeable difference between NNS-B and the other two groups for the /ʌ - ɛ/ continuum. Figure 5.5 (right column) shows the changes in sensitivity for each stimuli pair.

Table 5.7: /ɑ - ʌ/ continuum sensitivity (*d'*) scores, 2-step discrimination

	1-3	2-4	3-5
NNS-B	2.24	2.80	2.43
NNS-A	2.03	2.60	2.75
NS	2.78	3.12	3.19

The /ɑ - ʌ/ continuum yielded higher sensitivity scores than those obtained in the 1-step discrimination task (Table 5.7), but with no noticeable peaks except for the NNS-B group, which showed a peak at stimuli pair 2-4 (*d'*= 2.80). The NNS-A group showed a slightly raising pattern. The NS group also presented a raising pattern, but again with higher overall scores then the NNS groups. The Kruskal-Wallis test again showed no significant differences for any of the stimuli pairs.

Finally, the /ʌ - ɛ/ continuum triggered different responses among groups (Table 5.8). NNS-B showed overall considerably lower values and a mild peak for the stimuli

Table 5.8: /ʌ - ɛ/ continuum sensitivity (*d'*) scores, 2-step discrimination

	1-3	2-4	3-5
NNS-B	2.41	2.57	1.89
NNS-A	3.50	3.43	3.52
NS	3.11	3.73	3.60

pair 2-4 with a *d'* value of 2.80 with a decreasing slope afterwards; NNS-A shows no peaks. NS presented a different pattern, somewhat closer to the predicted pattern for categorical perception, with overall high values and a mild peak at the 2-4 token pair (*d'*= 3.73). The Kruskal-Wallis test showed just one significant difference: lower scores were obtained by the NNS-B group for the 3-5 stimuli pair (H (2) =8.09; $p < 0.05$).

5.5.4 Task 3: L1 labelling

Figure 5.6 shows the fitted probability values for the responses given by both groups (error bars at 95% confidence interval). In order to obtain these values, logistic regressions for both continua were performed. The best model for the /ʌ - ɛ/ continuum considers the interaction between stimuli and level as predictor, whereas the best fit for the /ɑ - ʌ/ continuum was a model with only stimulus as a predictor. This difference suggests that while proficiency affects perception of the /ʌ - ɛ/ continuum, it does not affect categorisations along the /ɑ - ʌ/ continuum.

NNS-B and NNS-A presented similar labelling patterns for both continua, with a clear mapping of /ɛ/ onto /e/, /ɑ/ onto /a/, and /ʌ/ mostly onto /a/ but also /o/. However, it is worth noting the dissimilarity in responses between NNS-B and NNS-A in the case of the /ʌ - ɛ/ continuum. On the other hand, the /ɑ -ʌ/ continuum triggered a strikingly similar categorisation pattern between the NNS groups with a very similar proportion of answers for all stimuli. The sole exception is stimulus 5 (=/ʌ/) where the categorisation counts for /o/ were slightly higher in the NNS-A group than in the NNS-B group, although the predictions given by the logistic regression were exactly the same for both groups.

5.5.4.1 NNS-B group

The /ʌ - ɛ/ continuum presented a very similar categorisation pattern to that of the L2 labelling task, which suggests that the L2 labels used in the previous identification task were, for the NNS-B subjects, exactly the same as those used here: the vowel in *cup* is the same as the one in *pan* (‘bread’) and the vowel in *bed* is the same as in *red* (‘net’). Table 5.9 shows the categorisation proportions for each stimulus.

In contrast, the /ɑ -ʌ/ continuum triggered a different behaviour. Here, the NNS-

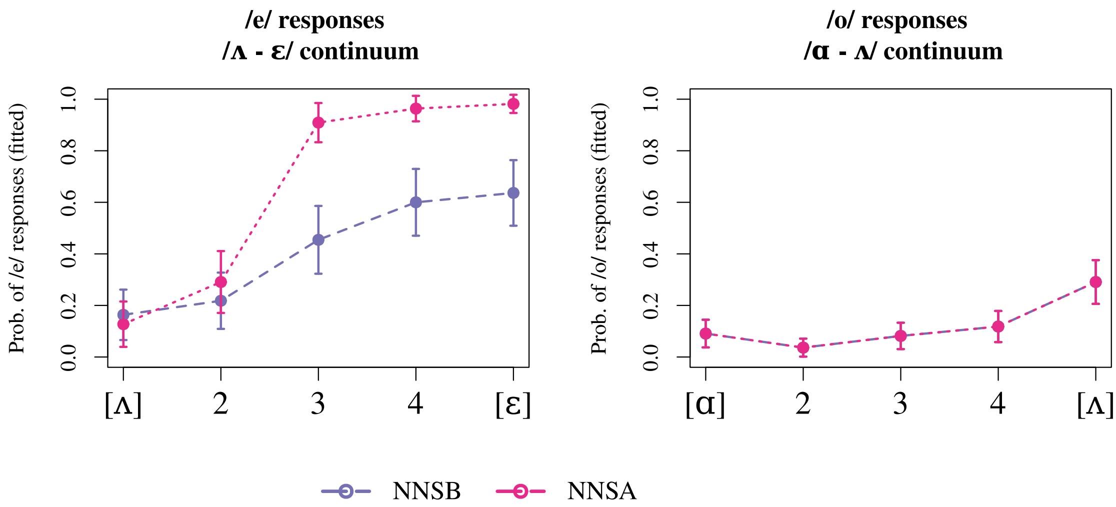


Figure 5.6: Results of L1 labelling task.

Table 5.9: Proportion of categorisations as L1 categories, /ʌ - ɛ/ continuum.

Stim	**1 (= ʌ)**		**2**		**3**		**4**		**5 (=ɛ)**	
Categ	a	e	a	e	a	e	a	e	a	e
NNS-B	0.84	0.16	0.78	0.22	0.55	0.45	0.4	0.6	0.36	0.63
NNS-A	0.89	0.11	0.73	0.27	0.11	0.89	0.02	0.97	0	1

B group has labelled the entire continuum as tokens of /a/, with a slight decrease in the counts as stimuli became acoustically closer to /ʌ/. The categorisation rates for /o/ do not reach 50% at any point along the continuum. However, it is worth noting that nevertheless there *are* counts of /o/ along the continuum, which increase in a linear way as stimuli come closer to /ʌ/. Table 5.10 (first row) shows the proportion of categorisations for each stimulus.

Table 5.10: Categorisation proportions, L1 labelling task. /ɑ - ʌ/ continuum.

Stim	**1 (=ɑ)**		**2**		**3**		**4**		**5 (=ʌ)**	
Categ	a	o	a	o	a	o	a	o	a	o
NNS-B	0.89	0.11	0.96	0.04	0.95	0.05	0.87	0.13	0.76	0.24
NNS-A	0.93	0.07	0.98	0.02	0.89	0.11	0.91	0.09	0.62	0.38

This shows that, regardless of /ʌ/ being chiefly categorised as /ɑ/, there is a 20% probability to be perceived as /o/, whereas tokens that are acoustically closer to /ɑ/ are almost certainly not being perceived as /o/. Ultimately, this suggests that /ʌ/ is a less prototypical exemplar of /a/ than is /ɑ/, a claim that can be confirmed by the L1 rating task (see section 5.5.5).

5.5.4.2 NNS-A group

Categorisation along the /ʌ - ɛ/ continuum by NNS-A subjects presented again a pattern strikingly similar to that of the L2 labelling exercise, with the sole difference that here the boundary lies exactly between stimuli 2 and 3. Just like the NNS-B group, this shows that for them the task was exactly the same, as both English labels mapped perfectly onto the L1 categories. The difference between the NNS-A and the NNS-B group regarding the level of certainty at endpoints remains the same, suggesting that NNS-A speakers consider vowels with more centralised values to also be good exemplars of their L1 vowel system. It is worth noting that NNS-A subjects have a large perceptual area for the /e/ category: tokens 3 to 5 were categorised as /e/ with a high level of certainty, regardless of their low and centralised values (token 3 of the /ʌ - ɛ/ continuum, with F1 = 832 Hz and F2 = 1589 Hz, has been categorised as /e/ in 80% of the counts).

Likewise, NNS-A, again, showed a perceptual pattern for the /ɑ-ʌ/ continuum that was very similar to that of the NNS-B group, with a slightly higher preference for the /o/ category when categorising stimuli that were acoustically closer to /ʌ/. The second row in Table 5.10 shows the corresponding categorisation proportions. Nevertheless, this preference is not enough to support a deflection hypothesis where /ʌ/ would be consistently mapped onto the second closest native category in the perceptual space. Should this be the case, the perceptual pattern would be characterised by a change in categorisations, with a much higher probability for categorisations of /ʌ/ as /o/ when reaching the relevant endpoint than the one observed here.

In sum, the L1 labelling task shows that, categorisation-wise, there is very little learning on the part of the NNS-A group. This learning can be mostly characterised as movement of the boundary zones, thus allowing for vowels with centralised values to be mapped onto the more peripheral categories of the L1. This higher tolerance for vowels with open-mid, centralised formant values can be observed in NNS-A when categorising tokens corresponding to /ʌ/: when the available labels are /a/ and /e/ (as in stimulus 1 of the /ʌ - ɛ/ continuum), then categorisation counts for /a/ reach ceiling. However, when the given labels are /a/ and /o/ (as in stimulus 5 of the /ɑ - ʌ/ continuum), then the counts are almost at chance level. NNS-B, on the other hand, are equally unsure as to whether a /ʌ/ is either /a/ or /e/, and even whether it is /a/ and /o/.

Given the high similarity between the two labelling tasks, it is worth probing whether L1 labelling is highly correlated to L2 labelling. A Chi-square test gave highly significant values for both groups of NNS (χ^2 = 218.73, p <0.001 for beginners, and χ^2 = 237.31, p <0.001 for advanced), which strongly suggests that categorisation of stimuli into L1 categories plays a role in categorisation of the same stimuli into L2 categories.

5.5.5 Task 3: L1 ratings

Figure 5.7 shows the ratings given to all tokens of each stimulus corresponding to the /ɑ - ʌ/ continuum; mean values are given in Table 5.11. Together with the high percentage of categorisations given to /a/, the NNS-B group (first row) also assigned higher ratings to the stimuli along the /ɑ - ʌ/ continuum that was categorised as /a/ (mean = 3.04) than to those stimuli that were categorised as /o/ (mean = 2.00). While a t-test comparing these ratings might not be meaningful given the differences in size (195 categorisations as /a/ versus only 30 as /o/), this same size difference corroborates the unlikeliness of rating any of these tokens as good exemplars of /o/. The NNS-A group (Figure 5.7, second row) again gave higher

Table 5.11: mean ratings given to stimuli as exemplars of L1 categories, /ɑ - ʌ/ continuum.

Stim	**1 (=ɑ)**		**2**		**3**		**4**		**5 (=ʌ)**	
Categ	a	o	a	o	a	o	a	o	a	o
NNS-B	3.14	3.00	3.06	0	2.83	3.50	3.27	1.00	2.90	1.71
NNS-A	3.71	3.33	3.75	3.00	3.58	2.60	3.46	2.50	3.57	2.41

ratings to the stimuli in the /ɑ - ʌ/ continuum than NNS-B (mean for categorisations as /a/ = 3.62; mean for /o/ = 2.56). However, the proportion in categorisations was similar between groups (244 categorisations as /a/ versus 31 as /o/). While a t-test comparing ratings within groups is again not possible given the sample size difference, a comparison between the two NNS groups shows that the ratings of stimuli categorised as /a/ were significantly higher in the NNS-A group (t(436.35) = 4.88, p < 0.001).

Stimuli corresponding to the /ʌ -ɛ/ continuum were rated by the NNS-B group mostly as poor exemplars of either one of the categories. In this case, the sample sizes were similar and consequently a t-test was performed. The comparison of the means of tokens categorised as /a/ versus those categorised as /e/ did not yield significant values (t(219.71) = -1.45, p = 0.14), suggesting that these subjects considered the stimuli to be equally poor exemplars of either /a/ or /e/. Table 5.12 shows the mean ratings given to each stimulus along the /ʌ - ɛ/ continuum, according to the way in which they were categorised as L1 vowels.

Table 5.12: Mean ratings for /ʌ - ɛ/ continuum stimuli, as exemplars of L1 categories /a/ and /e/.

Stim	**1 (=ʌ)**		**2**		**3**		**4**		**5 (=ɛ)**	
Categ	a	e	a	e	a	e	a	e	a	e
NNS-B	2.89	1.33	2.51	2.42	2.53	3.20	2.45	2.97	2.65	3.17
NNS-A	2.98	2.60	3.27	2.25	3.60	3.13	3.00	3.43	0	3.58

However, the NNS-A group gave better ratings to the stimuli corresponding to both categories. Again, the t-test yielded a rather high p-value (t (144.29) = 1.07,

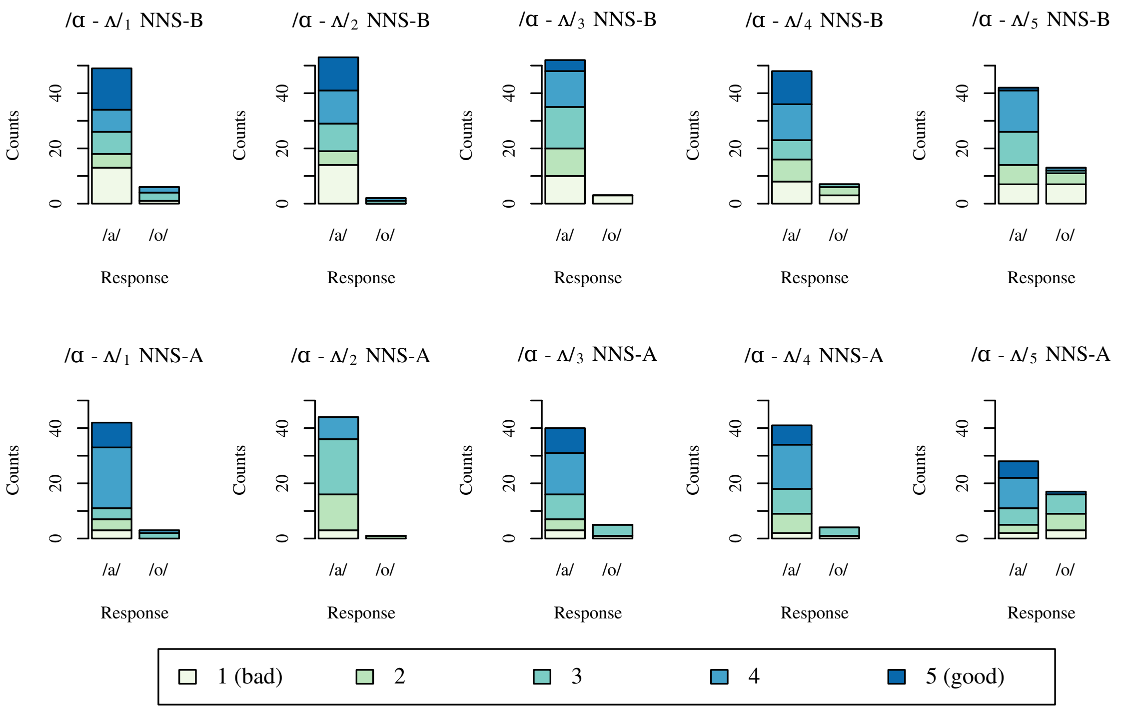


Figure 5.7: L1 rating task rasults by stimulus, /ɑ - ʌ/ continuum (/ɑ - ʌ/$_1$ = first continuum stimulus, i.e. /ɑ/).

p = 0.29), suggesting that subjects of this group neither considered the stimuli to be categorised as /a/ better than those categorised as /e/, nor vice versa.

A final comparison between groups confirmed again that NNS-A give higher ratings than NNS-B: the t-test between ratings of stimuli categorised as /e/ by both NNS groups ($t(204.81) = 2.12$, $p < 0.05$, one-sided), is only significant for $\alpha = 0.05$, but not for $\alpha = 0.01$. On the other hand, the comparison of the ratings given by both groups to stimuli categorised as /a/ yielded a more significant p-value ($t(155.92) = 2.42$, $p < 0.01$, one-sided).

5.5.6 Comparing L1 prototypicality ratings

Here I draw a comparison of the L1 prototypicality ratings assigned by each group and focus on categorisations that obtained more than 50%. Figures 5.9 and 5.10 show the stimuli identified as categories that obtained more than 50% of the labelling task; for instance, if stimulus 1 of the /ɑ - ʌ/ continuum was categorised as /a/ by a NNS group 70% of the time and as /o/ 30% of the time, then it will be a token mostly categorised as /a/. The figures show the stimuli as circles whose diameter equals the mean rating obtained multiplied by the percentage of highest categorisations. The size of each token is then a correlate of the strength with which the token is attracted by the category.

Figure 5.9a shows that the NNS-B group gave lower ratings to the tokens along

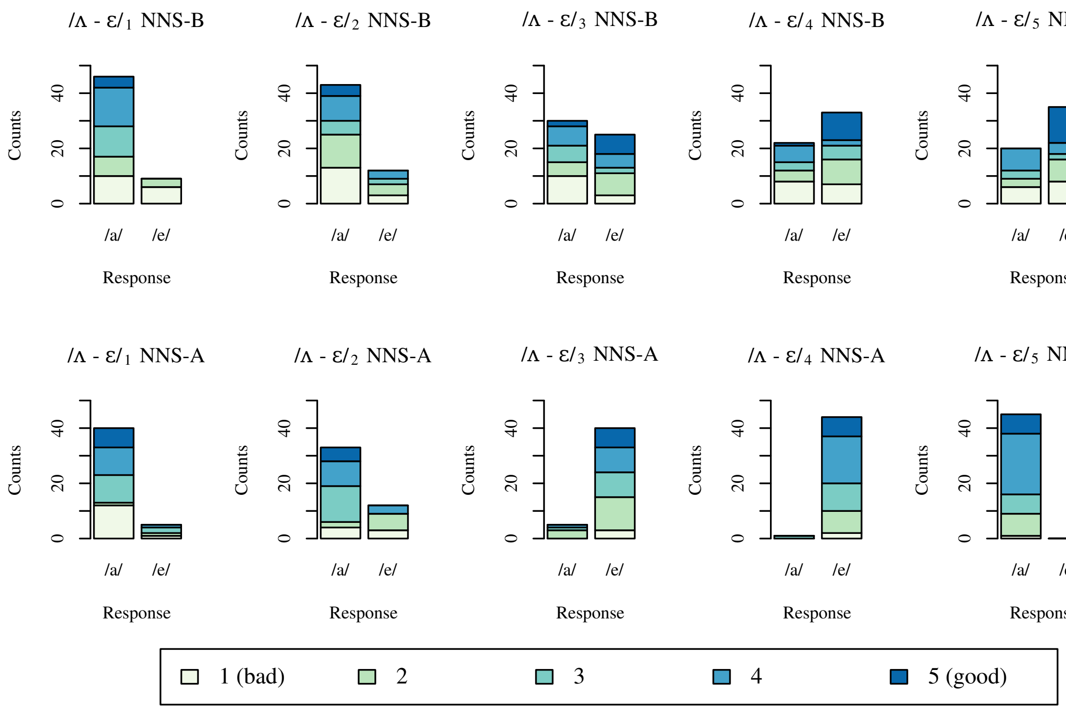


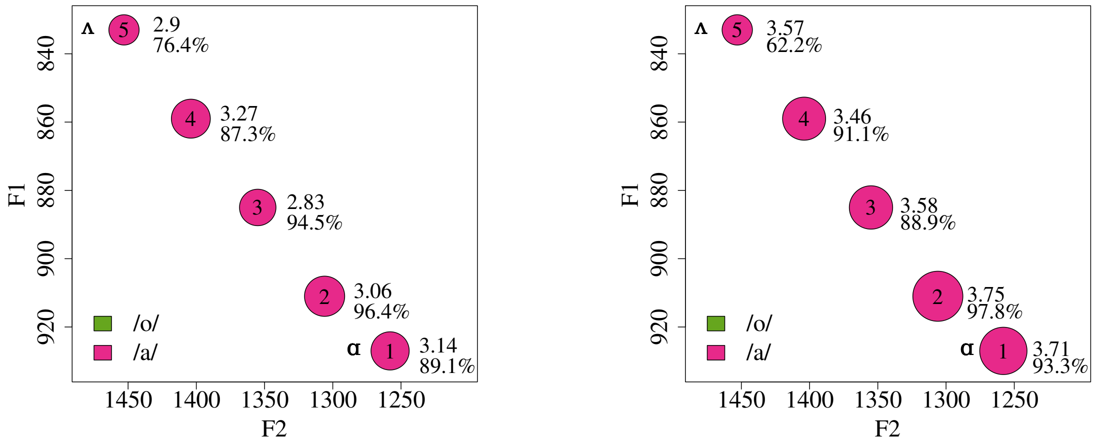
Figure 5.8: L1 rating task rasults by token, /ʌ - ɛ/ continuum ($/ʌ - ɛ/_1$ = first continuum token, i.e. /ʌ/).

the /ɑ - ʌ/ continuum, with nevertheless a high percentage of categorisations of them as /a/. It is worth noting the relatively little difference in rating means per token, which range between 2.83 and 3.27.

The NNS-A group (Figure 5.9b) assigned higher ratings to the /ɑ - ʌ/ continuum stimuli than the NNS-B group, although the range is, again, small (3.46 to 3.75). If the size of stimulus 5 (=/ʌ/, upper left) in both groups of NNS is compared, a trade-off between categorisation and prototypicality can be noted: both tokens are roughly the same in size, but the high ratings given by the NNS-A group compensate for the fewer categorisations.

The configuration of the perceptual space looks entirely different in NS (Figure 5.10). Firstly, two categories can be clearly seen, and secondly, the range of mean ratings is much higher: from 2.41 to 3.24. While the /ɑ - ʌ/, NNS-A ratings are not particularly high, it is worth noting that this higher range in ratings generates a noticeable difference in circle sizes, and that the smallest stimuli correspond to those that are at the category boundary (3 and 4). Hence, and just as expected, prototypicality increases at the endpoints, and decreases at the boundary.

Regarding the /ʌ - ɛ/ continuum, it is worth noting that while the labelling pattern was roughly similar among all groups (as shown in sections 5.5.1 and 5.5.4), ratings have different patterns. The /ʌ -ɛ/ continuum triggered a similar rating pattern to that described above: while the NNS-B group gave overall lower scores to the tokens in the L1 task (Figure 5.11a), NNS-A gave higher ones (Figure 5.11b).



(a) NNS-B

(b) NNS-A

Figure 5.9: Ratings and categorisation for /ɑ - ʌ/, nonnative speakers. While (a) shows smaller tokens than (b), both groups have categorised them mostly as instances of /a/.

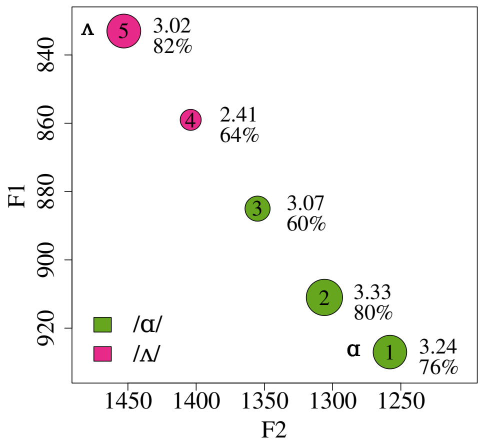
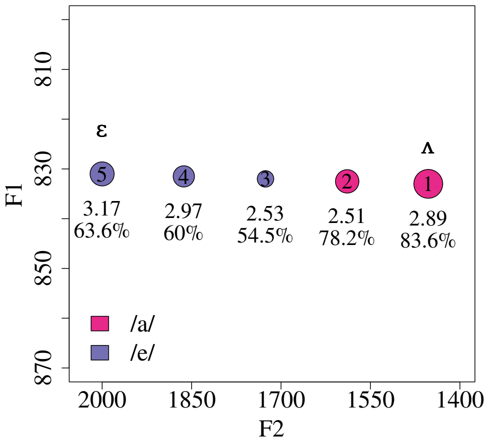


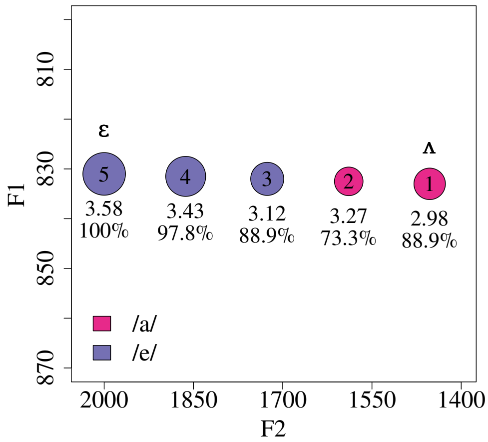
Figure 5.10: /ɑ - ʌ/ ratings and categorisation, native speakers. Tokens 1-3 are categorised as /ɑ/; tokens 4-5 as /ʌ/. Ratings are higher at the endpoints.

Importantly, both groups assigned higher scores to the tokens that were labelled as /e/, which are exactly the same ones that were labelled as /ɛ/ in the L2 task. Here NNS-B subjects have also assigned lower scores to stimuli 2 and 3, suggesting that the endpoints are somewhat more prototypical that these two stimuli, though not to a large extent given the overall low scores. NNS-A did not consider the tokens categorised as /a/ very prototypical: their scores were still higher than those given by the NNS-B group and there is a slight difference in scores between these and those that were categorised as /e/ (3.13 versus 3.38, respectively).

In sum, and regarding the way in which subjects assign ratings to the /ɑ - ʌ/ continuum stimuli, the NNS-A group seems to show a pattern similar to that of NNS-B, in that both groups *decrease* their ratings as the stimuli are acoustically closer to /ʌ/. However, even the lowest mean rating (stimulus 4, mean = 3.46) is



(a) NNS-B



(b) NNS-A

Figure 5.11: /ʌ - ɛ/ ratings and categorisation, nonnative speakers. While the categorisation patterns are the same in both groups, the ratings are lower in (a).

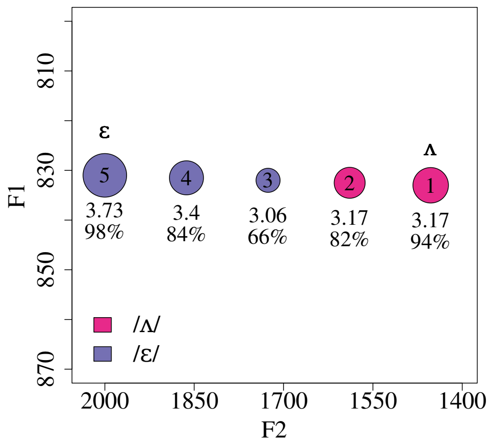


Figure 5.12: /ʌ - ɛ/ ratings and categorisation, native speakers.

higher than the best-rated stimulus by NNS-B (stimulus 1, mean = 3.14). NNS-A subjects then, accept stimuli that were previously identified as /a/ as much better exemplars of the category than NNS-B do.

On the other hand, the NNS-A group shows a rating behaviour that is not similar to either one of the other groups. Again, their ratings are overall higher than NNS-B, but unlike NNS-B and NS, their ratings do not increase as stimuli become closer to the endpoints of the continuum; rather, they decrease as they move from /ʌ/ to /ɛ/.

5.6 Conclusions

5.6.1 Changes according to proficiency

While the results of this experiment show us differences in perception between the NNS-B and NNS-A groups, such differences do not seem to indicate that experienced learners of L2 English have a new set of representations that includes a /ɑ/ - /ʌ/ contrast at the phonemic level. However, certain differences in discrimination and ratings suggest that the perceptual space is being reaccommodated given the new perceptual input. This ultimately suggests that intake does take place in NNS-A subjects, although not at a phonemic level because the system of representations remains the same as in the initial state; rather, their perceptual strategies for L2 speech perception include a different way of dealing with the acoustic input at a phonetic level.

5.6.1.1 L2 labelling and rating task: further evidence for a 2-to-1 mapping

The L2 labelling task showed that both groups of nonnative speakers have a similar system of representations: no new elements seem to have been added to the inventory, and there are no traces of a height distinction that could allow the creation of a new open-mid vowel category. This lack of distinction entails that the lower-back part of the vowel space is being populated only by a supercategory that takes values corresponding to both /ɑ/ and /ʌ/. At this point, it was observed that the /ʌ/ label was preferred by both NNS groups, though still with a considerable count of /ɑ/ categorisations. However, and aside from the preference in labels, both groups of nonnative speakers map all the /ɑ - ʌ/ continuum stimuli onto the same representations, in a similar manner. These labelling patterns can be compared to those obtained when labelling tokens of the /ʌ - ɛ/ continuum, where the lines for category counts clearly cross thus showing the presence of two different categories.

An important observation that must be made is that the representations used by nonnative speakers, when listening to L2 sounds, are not the result of a 2-to-2 vowel mapping, where /ʌ/ maps onto /a/ and /ɑ/ maps onto /o/ or vice versa. While this was certainly a possibility, given the orthographic <o> in *pot*, this was not the case; nonnative speakers would have shown the same categorical pattern in labelling observed in the /ʌ - ɛ/ continuum, if it were. Similarly, if speakers have associated the word *pot* with a rounded sound despite having heard the audio where *pot* is articulated as /pɑt/ and were therefore expecting a rounded sound (e.g./ɔ/) as a better candidate for the /ɑ/ label, then the counts of categorizations as /ʌ/ would have reached ceiling all along the continuum.

The /ʌ - ɛ/ continuum showed that whenever two L2 vowels are mapped onto two L1 vowels in a Two-Category assimilation, then the pattern in categorisation changes dramatically. Here, both ends of the continuum are being identified as different categories by both groups of nonnative speakers, which causes the crossed categorisation pattern: high counts of /ɛ/ and low counts of /ʌ/ on the /ɛ/ endpoint, and viceversa. It is worth noting, however, that the NNS-B group did not assign a proportion of counts as high as the other groups did to the endpoints.

The high categorisation counts for /ʌ/ all along the /ɑ - ʌ/ continuum were backed by the high prototypicality ratings given by the NNS-A group, which shows that (a) they consider /ɑ/, /ʌ/, and the other three stimuli between them as good exemplars of a category that they identify as /ʌ/; and that (b) all the stimuli along the continuum are good exemplars of /ʌ/. On the other hand, the fact that ratings for the /ʌ - ɛ/ continuum stimuli were higher for stimuli categorised as /ɛ/ shows that /ɛ/ is a much better fit for the /e/ category than /ʌ/ is for /a/; ultimately, this suggests that /ɛ/ is considered a clearer category than /ʌ/. However, it is worth noting here that the task itself (choosing between two categories) might have an effect in ratings.

Hence, the evidence gathered by the identification task suggests that more experienced non-native speakers have not created a target-like representation for these two vowel sounds; therefore, the representational task mentioned by the L2LP model did not take place. While this task in particular does not provide evidence on whether subjects perceive non-phonemic differences along the continuum, it shows that the entire /ɑ - ʌ/ continuum is being categorised in the same manner. Furthermore, the labels proved to be redundant, as they are not associated with different sounds. Instead, both labels were triggered in a 37-63% proportion in inexperienced learners, and 32-68% in more experienced learners.

L2 ratings show that, despite sharing the same system of representations, experience in the L2 affects judgement of the acoustic stimuli in terms of their goodness of fit: the more experienced in the L2 nonnative subjects are, the more likely it is that they will consider a certain stimulus to be a good member of a given category. However, these results should be interpreted carefully, since it is unlikely that the NNS subjects were rating the stimuli according to their goodness of fit to an L2 category; rather, the results obtained in the other tasks suggest that they evaluated the tokens to be members of an L1 category.

5.6.1.2 Discrimination: are there peaks in category boundaries?

Discrimination of adjacent tokens in a vowel continuum triggered patterns that do not correspond to the typical categorical behaviour seen in the categorical perception literature. Furthermore, all groups failed to show such pattern, for which the hypotheses made at the beginning of the chapter regarding a peak in discrimination where identification sets the boundary were not borne out. Since both groups of nonnative speakers were unable to set a category boundary in the labelling tasks for the /ɑ - ʌ/ continuum, peaks and overall changes in sensitivity were not expected; however, the sudden change in L2 categorisations between stimuli 1 (=/ɑ/) and 2 given by the NNS-B is reflected in the *d'* score raise when discriminating between stimuli 1 and 2. Likewise, the short, yet abrupt switch in labelling from *cup* to *bed* in the /ʌ - ɛ/ continuum can also be seen in the raise in discrimination between stimuli 2 and 3. In other words, and regarding nonnative speakers' discrimination behaviour, *d'* scores were expected to reinforce the results of a categorisation task where changes in categorisation along the continuum could be seen, but given the lack of these, the task was not able to provide information on that aspect.

Hence, the categorical perception phenomenon observed by Liberman et al. (1957) in consonants is absent here, due to the absence of categorical labelling patterns in nonnative speakers. However, a study by Fry et al. (1962) shows that both labelling and discrimination with synthetic vowel stimuli does not show the same sharp effects that consonants produce; the claim regarding differences in vowels was supported by Schouten et al. (2003). Importantly, native speakers also failed to display the categorical perception pattern in this experiment, where, instead of a peak, there is only a drop in *d'* when comparing stimuli 4 and 5; in the /ɑ - ʌ/ continuum both of which were mostly perceived as instances of /ʌ/.

On the other hand, the 2-step discrimination task yielded higher sensitivity scores for each stimuli pair, and mild peaks could be seen in some cases.

It is worth noting that the acoustic distance between the endpoints of the /ɑ - ʌ/ continuum is not particularly large: only ΔF1= 104 Hz and ΔF2= 195 Hz. Endpoints of the /ʌ-ɛ/ continuum, while having almost exactly the same F1, have a much larger frontness difference (ΔF2= 547 Hz). Hence, Euclidean distance between categories might be a good explanation for the low discrimination scores in the 1-step /ɑ - ʌ/ continuum and the similarities across groups.

The /ʌ - ɛ/ continuum showed that the similarity in labelling patterns have no direct relationship with sensitivity: all speakers have shown remarkably different sensitivity to stimuli pairs, where the pattern shown by each group cannot be explained on the grounds of labelling. However, the lower scores obtained by the NNS-B group in the 2-step task suggests very low reliance on F2 that the NNS-A does not display.

5.6.1.3 L1 labelling and rating: how different are IL representations from L1 phonemes?

L1 labelling showed that both groups of nonnative speakers of English are likely to perceive stimuli corresponding to the categories /ɑ/ and /ʌ/ as instances of the Spanish category /a/. While the second candidate for these stimuli is the Spanish category /o/, this was only likely to be triggered by the last /ɑ - ʌ/ continuum stimulus, namely /ʌ/. This ultimately suggests that these subjects' parameters of height are determined in a different way than those of native speakers and that backness is a less important perceptual cue; otherwise, the stimuli closer to /ɑ/ would have been mapped onto /o/.

L1 ratings show that NNS-A speakers have IL categories that are very similar to their original L1 categories, but seem to allow more deviant tokens of them. In this regard, L2 speakers of English have shown that they map the L2 categories /ɑ - ʌ/ onto L1-like IL categories; while these categories are not nominally different to those of the L1, their range of acceptable acoustic values has extended beyond the limits given by the L1. This can be seen in both continua, where NNS-A provided higher ratings to the given stimuli than subjects in the NNS-B group.

Labelling does not seem to have any relation to discrimination: labelling patterns were different along a given continuum by group and discrimination results were, in most cases, exactly the same. Euclidean distances are better predictors of discrimination, but they have no effect in labelling.

The ratings given by nonnative speakers to the stimuli based on their chosen labels, however, are significantly different: the NNS-A group gave higher ratings to each one of the stimuli. This ultimately suggests that their L1 representations are more tolerant to dispersion across the acoustic space. The NNS-A group has *heard* similar sounds, and they know that they are possible in the L2, but the labelling shows that they are unable to create two different categories for these sounds. On the other hand, the NNS-B group, with less experience input-wise, gives lower ratings to these vowels as they are not good exemplars of their categories, which we assume, at this point, are more likely to be the same as those of the L1.

5.6.2 Phonological intake in late-learners

The labelling tasks showed that nonnative speakers of English with Spanish as L1 do not *create* new perceptual categories that resemble those of native speakers of English. However, the differences in ratings and the overall similar discrimination behaviour across groups suggest that these learners are modifying their existent perceptual categories in terms of the range of possible tokens of them: however, these categories keep the same labels as those of the L1. This learning relates to the phonetic domain, as it entails storing certain phonetic detail in their long-term memory so that their categories accept these tokens as possible and frequent.

It seems then, that creation of a new perceptual category in L2 after a certain age may require more than just acoustic input. While specific training in perception of the relevant acoustic cues is possible (see McGuire, 2007; Logan, Lively, & Pisoni, 1991; Iverson & Evans, 2009, to mention just a few), the lack thereof does not produce the same results. While the evidence gathered so far points to the inability to create new categories, it does not imply that it is impossible; it only suggests that

(a) even higher amounts of acoustic input may be needed; and (b) that if a theory of second language acquisition where L2 speakers are able to learn a language in a target-like manner is taken into account, then these learners might have to resort to other cognitive mechanisms in order to achieve such a degree of performance; much more in line with the main tenets of the Fundamental Differences hypothesis.

In sum, the findings of this experiment suggest that the representations created by experienced learners of L2 English with Spanish as L1 are highly likely to be *nominally* the same as those in the L1, though with different boundaries and more tolerance to peripheral acoustic values. Labelling has shown that in cases where the initial state can be described as one of Single-Category assimilation (as in the /ɑ/ - /ʌ/ distinction); a set of two L2-like labels would generate chance-like categorisations in nonnative speakers along a continuum between these two vowels, with little to no variation from stimulus to stimulus. A different trend is seen in the L1 labelling task, in which the end of the /ɑ - ʌ/ continuum is perceived by nonnative speakers both as /a/ and the native category /o/. This shows that nonnative speakers are ultimately responding to the changes in vowel height, but that such height boundary is set further away than in the case of native speakers. However, the discrimination task has shown that more experienced speakers share a similar level of sensitivity with native speakers, both in 1-step and 2-step discrimination; thus, in terms of perception of differences, these two groups behave similarly. Finally, ratings have shown the effect of linguistic experience on what can be considered good exemplars of L1 categories.

5.7 Summary

This chapter presented Experiment 2 which aimed to investigate the phonological nature of IL representations in nonnative speakers of English with Spanish as L1. Since Experiment 1 provided evidence for a change in perception of differences between endpoints of a continuum and each one of the continuum points, further experimentation was needed in order to understand the way in which the system of nonnative representations caused these differences.

The chapter presented a rationale for each of the tasks used: labelling and rating with L1- and L2-like labels, and discrimination. While the former aimed to find a way to describe the representations stored by nonnative speakers and whether there are changes according to increased experience in the L2, the latter looked into nonnative speakers' ability to perceive differences in comparison with native speakers.

The experiments provided empirical support for the hypothesis that IL representations created by L2 speakers of English in Two-Category assimilation cases, such as the /ɑ/ - /ʌ/ distinction, are not phonemic insofar as these vowel sounds –

that would create minimal pairs in the L2 – are identified as members of the same category. However, nonnative speakers are nevertheless able to not only perceive acoustic differences in tokens of the same category – much in the same way as native speakers do, they also seem to change their parameters for goodness of fit when assessing the prototypicality of L2 tokens drawn from a nonnative-to-nonnative vowel continuum.

CHAPTER

SIX

CONCLUSIONS

La experiencia es un billete de lotería comprado después del sorteo. No creo en ella.

Gabriela Mistral

6.1 Phonemic split or subcategorial learning?

In this work I provide evidence to support my initial claim regarding the nature of perceptual vowel categories created by nonnative speakers with a smaller vowel inventory, and more specifically, the case of L1 Spanish late-learners of English. While my initial question refers to the phonological nature of the representations created in the IL when two vowels in the L2 undergo a Single-Category assimilation process, my claim is that such representations in late-learners are not of a phonemic nature; nevertheless, these representations are not exactly the same as those found in an initial state, where the L1 categories are being transferred with no modification. In fact, the system of representations does undergo changes, but they are not able to permeate the membrane of phonemic representations; rather, the existing categories have different boundaries than those originally set during L1 acquisition, and nonnative speakers become partially sensitive to sub-distributions of the tokens within the L1 acoustic space. I have claimed that a phonemic split of the initial L1 category into two different L2 categories is not possible, but that the size and shape of the representations transferred from the L1 change as a function of experience in the L2. Therefore, while /ʌ/ and /ɑ/ are both considered instances of the Spanish category /a/, nonnative speakers are able to discriminate between them, provided that the acoustic distance is large enough.

The experiments presented in the previous chapters suggest that late-learners of L2 English do not create new phonemic vowel categories for /ʌ/ and /ɑ/ as a function of high proficiency and/or experience. While Experiment 1 showed a departure from the initial state in terms of perception of differences in a discrimination task, Experiment 2 attempted to find a suitable type of label for the existing representations, and also check whether any other changes in perception within the phonetic domain could prove any type of boundary, should the only problem be that labels in the experimental design were not well chosen.

In fact, it could be argued that the results for the labelling tasks were unable to show category boundaries because the labels used were either not informative to the subjects (i.e. L2-like labels), or they were too large in terms of the perceptual space that they encompass (i.e. L1 labels). This, then, is when discrimination tasks play their main role. Since discrimination should show a peak at the category boundaries that were set by the labelling task, then nonnative speakers (who did not set boundaries in the labelling task) would not show any peaks. However, if the lack of boundaries was caused by the ineffectiveness of the labels and advanced learners do have TL-like categories but with different labels, then they ought to show a consistent ability to discriminate in the same way native speakers would.

It turns out that the target-like labels provided were not effective for purposes of identifying the differences in the incoming stimuli; in this regard, L1-like labels were better able to encode the difference. However, it became evident that /o/ was not a suitable enough label for /ʌ/, given that only around 40% of the total categorisations went towards /o/. Furthermore, the discrimination tasks in both experiments have shown that nonnative speakers are able to perceive differences in a manner similar to that of native speakers. It seems, then, that nonnative speakers are in fact able to *learn* something. But what is being learned? Apparently, they are able to notice phonetic detail in the L2 that their L1 does not require: differences *within* categories. I have referred to this in Chapter 3 as a **subsetting** scenario, or in terms of the type of learning with regard to the existent categories, *subcategorial learning*.

However, I have also claimed that the above represents a logical problem. The PME/NLM states that the perceptual space is warped around the prototypes of phonemic categories, which correspond to those of the L1; therefore, L2 tokens that fall within the range of attraction of these L1 'magnets' (and particularly more so in late learners) will then be attracted by these prototypes, for which learners will identify them as members of an L1 category. As I have already mentioned in Chapter 3, such phenomenon has been widely described in L2 phonology literature and the most relevant L2 phonology models such as PAM and SLM have described them. Likewise, categorical perception has shown that discrimination of two vowel sounds at distance X from each other *within* phonemic boundaries is reduced, and that it increases when two vowel sounds at the same distance X from each other are *on different sides* of the phonemic boundary. How is it possible, then, that nonnative speakers become able to perceive differences within categories? This specific type of acquired phonological knowledge is at the core of the concept of IL phonology: the L2 speaker's ability to create a category that corresponds nominally to those in the L1, but whose phonetic substance allows for subsetting and full inclusion of otherwise borderline acceptable tokens.

6.2 The nature of IL representations

Experiment 1 demonstrated two important aspects of nonnative speech perception. Firstly, it showed that nonnative speakers are unlikely to perceive phonetic differences between /ɑ/ and /a/, thus suggesting that the former is mapped onto the latter. Likewise, /ʌ/ was also perceived as the same when discriminating against /a/. This is the first piece of evidence towards the main assumption that the /ɑ/ - /ʌ/ contrast is a case of what the PAM model (Best, 1995) refers to as Single-Category assimilation, where two L2 sound categories are mapped onto one L1 representation. Secondly, Experiment 1 has also shown that the predictions of the PAM model were borne out. PAM predicts that discrimination between two sounds that are being perceptually assimilated onto the same L1 category will be poor; a discrimination task along the /ɑ - ʌ/ continuum showed that this was exacly the case for inexperienced speakers of L2 English.

Experiment 1 also showed that there were differences among groups of subjects. And crucially, the experiment located the group of experienced speakers of L2 English (NNS-A) on a step between less experienced L2 speakers and native speakers of English. While the less experienced L2 speakers were unable to perceive differences when comparing continuum stimuli against endpoints, native speakers reached ceiling performance; NNS-A subjects, on the other hand, reached an intermediate state where perception of phonetic differences along the continuum was higher than chance, but less than ceiling.

This intermediate pattern has two implications for a theory of IL representations. First, it proves that experience may lead to enhanced perception of differences in the acoustic input, thus ruling out the possibility of an absolute no-learning scenario: more exposure to the L2 is able to trigger changes in perception. And second, this change in perception does not necessarily lead to clear-cut, categorical distinctions between sounds; rather, behaviour seems to be probabilistic, where the probability that a certain difference will be heard is high, but not as high as in native speakers of English.

All in all, Experiment 1 suggested that learning had taken place in NNS-A subjects, and that this learning is not deterministic. At this point it seems evident that a perceptual task is taking place; however, it remains to be seen whether this type of L2 speaker has also carried out a representational task where the two sounds in question can be identified as two different categories.

In Experiment 2, the L2-labelling task showed that late-learners with advanced knowledge of English (NNS-A) were unable to identify the endpoints of the /ɑ - ʌ/ continuum as different categories. Furthermore, this categorisation pattern was not different from that of speakers with little knowledge of English; this evidence does not support the hypothesis that more experience in the language leads to creation of new perceptual categories. This experiment showed that there was a preference for choosing /ʌ/ along the continuum, which suggests that this label is closer to the acoustic values stored in the speakers' perceptual representations. Likewise, this experiment shows that there is no Two-Category assimilation of the English vowels /ɑ/ and /ʌ/ onto the Spanish vowels /a/ and /o/; should this be the case, then the NNS-A group would have displayed the same crossing-lines pattern as that triggered by the /ʌ - ɛ/ continuum.

Ratings of the stimuli corresponding to the /ɑ - ʌ/ continuum show that inexperienced speakers (who gave low ratings to all tokens) do not consider these continuum stimuli as prototypes of any of the perceptual categories that were proposed by the forced-choice task; in other words, these inexperienced nonnative speakers of English judged the heard stimuli as poor instances of a category that was slightly better represented by /ʌ/ than /ɑ/. This seems to be a direct consequence of the fact that less experience in the L2 makes the listener less able to judge the sound listened to as something that has already been heard before, and therefore recognise it as a frequent/acceptable token of a category. On the other hand, advanced speakers do seem to recognise these stimuli as good exemplars of a category, which as the labelling suggests, is a supercategory that allows a wide range of acoustic values; such a supercategory can be recognised as the L1 phonemic category /a/, but the differences in boundary zones and the way the expected distributions are different from the initial-state /a/ category.

Discrimination was the least informative of all the three perceptual tasks, probably due to the relatively short acoustic distance between /ɑ/ and /ʌ/. However, the fact that no group-specific perceptual patterns were observed suggests that the task tapped more phonetic instead of phonological knowledge, which ultimately points to the fact that the same level of performance can be expected from all groups, therefore sustaining the hypothesis that linguistic experience does not affect perception of acoustic differences. These conclusions are in line with Werker and Tees (1984), who stated that non-targetlike perception in nonnative speakers cannot be attributed to sensorineural loss and are instead a result of linguistic experience.

Finally, the L1 labelling task provided further evidence that L1 Spanish speakers of English map both English vowels /ɑ/ and /ʌ/ onto the L1 category /a/. While the L2 labelling task shows that either /ɑ/ or /ʌ/ are suitable labels for the /ɑ - ʌ/ stimuli, this labelling task shows that these stimuli are perceived as members of the /a/ category, with the exception of tokens corresponding to the vowel /ʌ/, which given the similar proportion of categorisations of it as either /a/ or /o/, seems to be at the boundary between these two vowels.

In this regard, ratings showed that more experienced speakers consider the tokens of the /ɑ - ʌ/ continuum as good exemplars of /a/. The fact that for these speakers this category is the best fit for the stimuli indicates that their /a/ category has expanded, therefore accepting less L1-like tokens as still good instances of this category.

Furthermore, when comparing the labelling patterns of advanced speakers in both L1 and L2 tasks, an important paradox unfolds: advanced speakers prefer the label /ʌ/ to categorise stimuli along the entire /ɑ - ʌ/ continuum, but when confronted with a prototypical token of /ʌ/ in the L1 labelling task the counts of categorisations as /o/ reached 40%. This is perhaps the most compelling evidence that advanced speakers are sensitive to phonetic differences, but they have not developed an L2 label to encode these differences; rather, they make use of their L1 categories, which are by definition limiting in the sense that they are too 'wide' to encode this level of phonetic detail.

These findings lead us to conclude that advanced speakers of English have not created two new different perceptual categories for the vowels /ɑ/ and /ʌ/, or at least not at the phonemic level. However, it seems clear at this point that these speakers have undergone learning which, although not phonemic, represents a qualitative change in their perceptual grammar (see below). What late-learners can do, however, and only by means of input, is to perceive phonetic detail in certain tasks, in a manner similar to that of NS. This happens after copious amounts of input have been received. Late-learners will then be able to store long-term memory representations of that phonetic detail, but these representations are coded as sub-members of a supercategory that corresponds to the L1 category onto which these L2 vowels were originally mapped. I claim that these long-term encodings of phonetic detail are the resulting representations of interlanguage phonology.

6.2.1 Specific changes in perception

From these experiments three specific changes in perception can be identified:

6.2.1.1 Subcategorial learning

Experiment 1 showed that perception of phonetic differences is available to advanced speakers, however, Experiment 2 suggested that these learners have not created a new system of representations. While the notion of perceiving differences *within* categories may sound counterintuitive from the classic structuralist viewpoint (and particularly so in vowel sounds), nonnative speakers are able to partition their L1 category /a/ into two different phonetic categories, where one end of the /ɑ - ʌ/

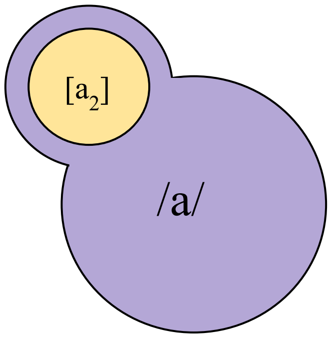


Figure 6.1: An IL category (in yellow). The L1 category /a/ holds two different distributions, with one of them, [a_2], created when listening to tokens of English /ʌ/. Differences can be noticed by listeners during discrimination tasks, but not identification.

continuum (the open-mid) does not map very well; this vowel is then pushed apart from the rest of the more prototypical category values and creates its own zone. This leads to a reshaping of the /a/ category, which can be visualised as in Figure 6.1.

6.2.1.2 L1 category enlargement

From the previous aspect it can be inferred that the category /a/ behaves as a supercategory in the L2 that takes values that are further away from a prototypical /a/ as acceptable tokens of it. This entails an enlargement of the perceptual L1 category, and it is the main effect of what was defined in Chapter 2 as the 'IL loop': a process whose output feedbacks itself in such a way that creation of a new category implies a bootstrapping process where the only possible mechanism to make this possible is noticing; however, noticing seems to be unavailable at all times and is better attested when nonnative speakers have a reference sound to discriminate against; identification, at least in isolation, did not trigger any effects that could be due to the presence of noticing.

Unfortunately, category enlargement may bring with it an effect that is opposite to the desired. If tokens of the L2 start fitting better into L1 categories, then the probability that learners split their category does not increase; rather, it decreases. Since noticing the phonetic differences, and more particularly, noticing them in such a way that they can be perceived as non-members of the L1 category is crucial in order to achieve category splitting, then the fact that these tokens become better fits for the L1 category makes the split impossible.

This is perhaps the most critical aspect of L2 speech perception: *once an L1 vowel, always an L1 vowel.* By this we mean that the constraint PERCEIVE in the L1 perceptual grammar is so highly ranked that as soon as one 'deviant' (i.e. L2) token is categorised as an L1 vowel, then the next L2 token will be again categorised as such; the grammar feeds itself so as to efficiently categorise tokens of the L2 into the existing representations. This mechanism, however, does not work to the detriment of perception of within-category differences in the specific context of a discrimination task.

6.2.1.3 Increased sensitivity to height distinctions

Perception of the stimuli corresponding to the /ɑ - ʌ/ continuum offers two different manners to divide the perceptual space: one where /ɑ/ is the less prototypical token, and another where /ʌ/ is less prototypical. The L1 categorisations showed that the latter was more ill-fitted as a member of /a/, which demonstrates that advanced speakers become more sensitive to the height distinction. In a different perceptual configuration /ɑ/ would have been considered a worse fit, given its backness; this could have made it more likely to be perceived as /o/. Likewise, it is worth noting that despite the centralised values of /ʌ/, its likelihood to be perceived as /o/ increases as a function of distance from /ʌ/.

6.2.1.4 Probabilistic categories

However, the advanced learners' ability to perceive phonetic differences is far from showing a deterministic behaviour. Instead, advanced learners seem to show a probabilistic behaviour in discrimination, which is higher than chance; however, this advantage over less experienced speakers is not enough to reach native-like performance. Sensitivity tests in Experiment 1 showed how this in-between sensitivity to the difference between /ɑ/ and /ʌ/ is given by a change in perceptual mechanisms that are nevertheless not robust enough to offer native-like reliability in discrimination.

6.3 Is there a CPH for phonology?

It is tempting to say that the data obtained with these experiments has implications for the CPH. However, two points must be taken into account. First, the highly proficient speakers of L2 English who participated in the experiments have spent different lengths of time living in an English-speaking country, which may have had an effect on their ability to create new perceptual categories. Nevertheless, and as mentioned in the section above, the 'IL loop' would make the amount of input an irrelevant factor, given that the supercategory grows as more input is categorised into it. Therefore, it is worth exploring whether even more input (e.g. when the years of consistent use of the L2 are more than those of using the L1) results in a category split by means of distributional learning.

Secondly, a crucial distinction must be made regarding the possible extent of the CPH in phonology. While a first approach could be more deterministic in terms of whether late-learners of an L2 can learn *anything at all*, a second (and more sensible) approach would focus on the aspects of the L2 phonology that are possible to acquire after the L1 has reached its final state. In this regard, and following the tenets of the L2LP model, the data obtained in these experiments suggest that *phonetic* learning (i.e. noticing differences in the acoustic signal) has been carried out by proficient learners, but that the *representational* task may either (a) take place after the amount of L2 input is large enough (how much?) or (b) not take place at all.

As an explanation for the lack of native-like perceptual representations of L2 segments that are not contrastive in the L1, I propose that the main reason is the loss of access to certain aspects of the language learning faculty, which ultimately affects segments with the aforementioned characteristics, in late-learners of an L2. More specifically, the ability to break down the perceptual space into *more* categories than those given by the L1, and being able to generate a prototype and a label for new categories in the same was as during L1 acquisition, seems to be lost; or at least it would need an unestimated amount of input for category creation to take place. Furthermore, and following Young-Scholten (1994) an important characteristic of L2 learning is that the Subset Principle (i.e. that in first language acquisition the learner chooses the most restrictive grammar out of two competing grammars) does not apply; hence, if the L2 is a superset of the L1, then the L2 learner will be unable to *notice* the target grammar. In our case study, we can understand the distinction between /ɑ/ and /ʌ/ as a more restrictive way to divide the acoustic space, whereas the L1 has just one perceptual category with a rather large space. Hence, the rules that divide the perceptual space of the L2 can be understood as a subset of those in the L1, insofar as the L1 has fewer boundaries than the L2. Under this assumption it can be stated that negative evidence (i.e. "a token with an F1 of 640 Hz is not /a/", or in lay terms, "what you think is an /a/ is actually a different vowel") is the only way to effectively break the sub-optimal perceptual loop that causes the perceptual /ɑ/ - /ʌ/ merger. Ths would in turn imply a highly specific intervention involving manipulation of perceptual cues, or drills with minimal pairs and corrective feedback.

6.4 UG access

Since the ability to create new labels for the perceptual categories that correspond to the L2 appears to be limited, there is a possibility that UG access in late-learners of an L2 is reduced. The division of the perceptual vowel space into further categories seems to be a particularly difficult task, with a height dimension that needs to be reconfigured from three (open, mid, and close) to four levels (open, open-mid, close-mid, and close). Likewise, if a feature-based approach to perception is assumed, then the emergence of a [tense] feature is expected; however, this seems to be again a difficult task since the tense/lax distinction is not present in Spanish and therefore it presents again the same bootstrapping problem that can be seen in creating a fourth height parameter. While acquisition of L2 categories by means of redeployment of L1 phonological knowledge has been attested (Archibald, 2005), this is again not the case with the /ɑ - ʌ/ distinction for native speakers of Spanish: no [tense] feature can be redeployed from any other pair of contrastive sounds in the L1.

It follows from the above that the type of learning involved in acquisition of the /ɑ - ʌ/ contrast has implications for a theory of learning in SLA. As mentioned in Chapter 2, UG-based approaches to SLA can be grouped in three different hypotheses: full access, partial access, and no access. Since the experiments reported in the previous two chapters have shown that input alone does not seem to be enough for the desired intake to take place, then the full access hypothesis needs to be discarded, or at least revised in terms of what will be considered as UG-driven acquisition and what can be better explained by considering other cognitive skills involved in the learning process. However, and as mentioned in the subsection above, this verdict may be overturned if target-like perception in late-learners can be attested; what remains to be determined is whether that amount of input, in a time span that allows for a myriad of other factors to play a role, can be still considered part of UG mechanisms in action, or if acquisition of these sounds is better described by taking into account the several aspects summarised by the Fundamental Differences hypothesis (Bley-Vroman, 1989). Furthermore, even if target-like *performance* in late-learners is attested, it is still unclear whether this could mean that target-like perceptual representations have been formed in the late-learner's grammar, or if it is an IL representation that performs with a reasonable degree of robustness.

Given the results of the experiments, however, it seems that the ability to create new perceptual representations is fairly limited after a certain stage of life, unless (perhaps) extensive training takes place; this possibility, it should be stressed, still differs from the UG tools at hand during L1 acquisition.

6.5 Implications for L2 phonology theory

6.5.1 On the explanatory adequacy of the current L2 phonology models

As discussed in Chapter 3, the current L2 phonology models have focused on different aspects of L2 phonology. While PAM focuses on naïve nonnative perception and the resulting perceptual assimilation processes, SLM focuses on ultimate attainment but in a more speculative manner; finally L2LP thoroughly describes the learning processes involved in the acquisition of perceptual categories, but does not make a specific claim about the nature of such categories.

All in all, these three models combined make specific contributions towards the research question posed in this thesis. Importantly, the results of the experiments suggest that PAM correctly predicts poor discrimination in a Single-Category assimilation scenario such as the /ɑ/ - /ʌ/ distinction by native speakers of Spanish. In this regard, these experiments also suggest that Single-Category assimilation represents a 'sand trap' for nonnative speakers as the probability of creating a new perceptual category in this condition is low.

In light of the experimental findings of this research, SLM seems to have the best approach to the question on the nature of IL representations by assuming that these are *phonetic categories*. However, the SLM claims are somewhat contradictory, as on one hand the framework firmly advocates against the CPH; nevertheless, by stating that these representations are of a phonetic nature the possibility that they are phonemes is, if not explicitly, at least partially denied. In this regard, a generative approach to L2 acquisition by means of full access to UG should then draw a line in terms of the type of representations that late-learners are able to acquire: there is full access *if and only if the created representations are phonemic*.

Finally, L2LP provides a very thorough conceptualisation of perceptual learning in L2, by drawing the distinction between perceptual and representational tasks and how these tasks are mediated by learning new perceptual cue weightings for the L2 input. However, the results obtained here did not provide evidence regarding the representational task; instead, it provided sufficient evidence that the perceptual task was taking place without generating new perceptual slots for emerging categories, therefore leading to the suboptimal subsetting outcome.

6.5.2 Acoustic and featural approaches

The experiments have also challenged the acoustic similarity approach to L2 speech perception in that /ʌ/, despite having the same acoustic values as /a/, was not judged as a prototypical exemplar of /a/; rather, experienced nonnative speakers gave /ʌ/ a considerable percentage of categorisations as /o/. In this regard, it was expected that /ʌ/ would be identified as /a/ to a much larger extent and that groups of nonnative speakers would give it high prototypicality ratings as a member of /a/. Hence, what was presented in Chapter 3 as the phonetic category [a_2] was not based on the acoustic values for /ɑ/; rather, it is a category that takes the acoustic values of /ʌ/.

This shows then, that while phonetic similarity plays a role in that there is some correlation between the phonetic values of L2 sounds and the category where these sounds will be mapped to, listeners also make use of abstract phonological features in order to decide how to divide the perceptual space. Nonnative speakers seem to perceive this contrast by assuming that [+low] is by definition *lower* than the actual production values for /a/. This might in turn call for an extension of L2LP that explicitly adds abstract phonological features, to which the acoustic values are mapped.

6.5.3 The L2LP model revisited

While the theoretical value of L2LP cannot be denied insofar as it provides a formal account of the L2 speech perception process throughout the different learning states, it fails to account for perceptual category creation: it cannot formally account for the appearance of a new perceptual candidate, should that be the case. Likewise, the model cannot specify the phonological nature of the new representation, which is the goal of this work. The subsetting outcome that has been proposed as an explanation to the results obtained in the experiments is then incompatible with the formalisations proposed by the L2LP model. In this respect, Chapter 3 proposed a set of auditory-to-phonological constraints (cue constraints), with a constraint family *New, which militates against creation of new categories. However, *New is not a different type of constraint; it is just a set of cue constraints that refer to L2-like categories, exactly as those found in the L2LP model. Demotion of this constraint, along with the appearance of *Warp constraint, can be seen as the only way to justify the presence of a new phonemic candidate.

On the other hand, the creation of a perceptual category that does not hold a phonemic status in the listener's grammar poses a significant problem regarding the type of constraints that can be used in order to prevent creation of new phonemic categories, but at the same time allow for the appearance of a phonetic category on the basis of noticing the mismatch between the prototypical values for the existing (L1) category and the new L2 category. In this work I have proposed that a model that attempts to account for such subphonemic phenomena in L2 must take into consideration not only auditory-to-phonological constraints (i.e. a perceptual cue that maps acoustic values onto phonological categories) but also auditory-to-auditory constraints, which allow for less distortion of the perceptual space. Thus, an auditory-to-auditory constraint that prevents categorising a value of 780 Hz into a vowel category whose prototypical acoustic values are of 640 Hz could allow for the creation of a perceptual candidate for these values, without warping them 100 Hz further away in the acoustic space but also without creating a new phonemic category.

However, the results presented in Chapter 5 show that perception of /ʌ/, despite having the same perceptual values for the Spanish category /a/, is not perceived as a prototype of the category, as mentioned in the subsection above. Hence, if we assume a category /a/ whose prototypical F1 values in terms of *perception* are around 780 Hz, then the *Warp constraint should then be relevant when listening to tokens of /ʌ/, and not /ɑ/, as it was presented in Chapter 3. Tableaux 6.1 and 6.2

show then the correct formalisation, with the *WARP constraint militating against mapping of an F1 of 640 Hz (=F1 of /ʌ/) onto /a/; Chapter 3 showed a *WARP constraint that prevented warping of an F1 of 780 Hz (=F1 of /ɑ/) to /a/.

Table 6.1: Perception of /ʌ/ in the subsetting outcome.

[640, 1350]	*NEW	640 Hz not /o/	*WARP ([640], /a/)	640 Hz not /a/
a. /a/			*!	*
b. ☞ [a_2]				*
c. /o/		*!		
d. /ʌ/	*!			

Table 6.2: Perception of /ɑ/ in the subsetting outcome.

[780, 1240]	*NEW	780 Hz not /o/	780 Hz not /a/	1240 Hz not /o/	1240 Hz not /a/
a. ☞ /a/			*		*
b. /o/		*!		*	
c. /ɑ/	*!				

Nevertheless, the problem of this model is that it still cannot account for the creation of new phonemic categories. If a *WARP constraint is assumed (which can be seen as a consequence of the noticing process), it is still unclear what causes demotion of a *CATEG constraint. This is one of the most important challenges that remain from this study, and that call for a better articulation of formal accounts of speech perception in the field of second language acquisition.

6.6 Future directions

While this work has provided evidence towards a null hypothesis regarding category creation in L2, it is evident that more research is needed in order to move towards a theory of L2 category creation in late-learners within a UG framework.

In the first place, the question whether categories can be created has been partially answered: it is clear that speakers who have been tested and qualified as highly proficient, and/or have spent a reasonable amount of time in an immersion setting, do not have these representations; in fact, a paradox between the increased ability to notice phonetic detail and the feedback loop created by repeated categorisations of tokens of the two L2 vowels in question as the same sound might after all be resolved in the way of category creation, after an undetermined amount of time. In this regard, a first step is to repeat these tasks on subjects who have spent larger amounts of time in an L2 immersion environment.

A second step is to test whether these results can be replicated with other vowel contrasts that are not present in the L1 (i.e. /i/ - /ɪ/ or /ɒ/ - /ɔ/), and see whether the representational task is easier in cases where the perceptual cues involved are different: F1 only, F2 only, duration. Furthermore, the /æ/ - /ɑ/ contrast would also provide further information regarding the validity of a feature-based approach to speech perception, a question that was already posed by Barrios et al. (2016), though the results suggest that feature availability does not result in contrast creation.

Finally, one further step is to compare these results with production data: are the vowels also merged in pronunciation? Are their targets deviant with regard to the L1 and also the L2? In the case that production data does not match that of perception (as is suggested by the results of Bohn & Flege, 1997), it is worth wondering whether target-like speech perception is the final frontier of SLA.

References

Antoniou, M., Tyler, M. D., & Best, C. (2012). Two ways to listen: Do L2-dominant bilinguals perceive stop voicing according to language mode? *Journal of Phonetics*, *40*(4), 582–594.

Apoussidou, D. (2007). *The learnability of metrical phonology.* Thesis (doctoral). Universiteit Amsterdam &LOT.

Archibald, J. (1993). *Language learnability and L2 phonology: The acquisition of metrical parameters.* Dordrecht: Kluwer Academic Publishers.

Archibald, J. (2005). Second language phonology as redeployment of L1 phonological knowledge. *Canadian Journal of Linguistics*, *59*, 285–314.

Barrios, S., Jiang, N., & Idsardi, W. J. (2016). Similarity in L2 phonology: Evidence from L1 Spanish late-learners' perception and lexical representation of English vowel contrasts. *Second Language Research*, *32*(3), 367–395.

Bassetti, B. (2008). Orthographic input and second language phonology. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 191–206). Clevedon, UK: Multilingual Matters.

Berwick, R. (1985). *The acquisition of syntactic knowledge.* The MIT Press.

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience : Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press. Retrieved from http://ci.nii.ac.jp/naid/10018033931/en/

Bialystok, E., & Hakuta, K. (1999). Confounded age: linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Mahwah, NY: Erlbaum.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second language acquisition. *Journal of Memory and Language*, *44*, 235–249.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning. *Linguistic perspectives on second language acquisition*, 41–72.

Boersma, P. (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, 43–58.

Boersma, P. (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives.* Holland Academic Graphics/IFOTT.

Boersma, P. (1999). On the need for a separate perception grammar. *Ms. Rutgers University: Rutgers Optimality Archive*, *358*.

Boersma, P. (2009). Cue constraints and their interactions in phonological perception and production. In P. Boersma & S. Hamann (Eds.), *Phonology in perception.* Mouton de Gruyter.

Boersma, P., & Hayes, B. (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry*, *32*(1), 45–86.

Boersma, P., & Weenink, D. (2012). *Praat: doing phonetics by computer (version 5.3.10) [computer program]. retrieved from http://www.praat.org/.*

Bohn, O.-S. (1995). Cross language speech production in adults: First language transfer doesn't tell it all. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 279–304). Timonium, MD: York Press.

Bohn, O.-S., & Flege, J. E. (1997). Perception and production of a new vowel category by second-language learners. In A. James & J. Leather (Eds.), *Second-language speech: Structure and process* (pp. 53–74). Berlin/New York: Mouton de Gruyter.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, *19*, 447–465.

Bradley, T. (2014). Fonología de laboratorio. In R. Núñez Cedeño, S. Colina, & T. Bradley (Eds.), *Fonología generativa contemporánea de la lengua española* (2nd ed., pp. 319–366). Washington D.C.: Georgetown University Press.

Bradlow, A. (1995). A comparative acoustic study of English and Spanish vowels. *Journal of the Acoustic Society of America*, *97*(3), 1916–1924.

Broselow, E., & Finer, D. (1991). Parameter setting in second language phonology and syntax. *Second Language Research*, *7*(1), 35–59.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*, 201–251.

Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, *14*, 136–193.

Carlisle, R. S. (1988). The effect of markedness on epenthesis in Spanish/English interlanguage phonology. *Issues and Developments in English and Applied Linguistics (IDEAL)*, *3*, 15–23.

Carroll, S. (2001). *Input and evidence: the raw material of second language acquisition.* Amsterdam: John Benjamins.

Christensen, R. H. B., & Brockhoff, P. B. (2009). Estimation and inference in the same-different test. *Food Quality and Preference*, *20*(7), 514 - 524. doi: https://doi.org/10.1016/j.foodqual.2009.05.005

Christensen, R. H. B., & Brockhoff, P. B. (2016). *sensR: An R-package for sensory discrimination.* (R package version 1.4-7. http://www.cran.r-project.org/package=sensR/)

Colantoni, L., Steele, J., & Escudero, P. (2015). *Second language speech.* Cambridge: Cambridge University Press.

Corder, S. P. (1967). The significance of learner's errors. *IRAL-International Review of Applied Linguistics in Language Teaching*, *5*(1-4), 161–170.

De Houwer, A. (2009). *Bilingual first language acquisition.* Bristol, UK: Multilingual Matters.

Dell, F. (1981). On the learnability of optional phonological rules. *Linguistic Enquiry*, *12*, 31–37.

Diehl, R. L., & Kluender, K. R. (1989). On the objects of speech perception. *Ecological Psychology*, *1*(2), 121-144.

Ellis, R. (1994). *The study of second language acquisition.* Oxford: Oxford University Press.

Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization.* Utrecht University & LOT.

Escudero, P. (2009). The linguistic perception of similar L2 sounds. In P. Boersma & S. Hamann (Eds.), *Phonology in perception* (pp. 152–190). Berlin: Mouton de Gruyter.

Escudero, P., Benders, T., & Lipski, S. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, *37*, 452–465.

Escudero, P., & Boersma, P. (2003). Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. *Proceedings of the 25th Annual Penn Linguistics Colloquium. Penn Working Papers in Linguistics*, *8*, 71–85.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551–585.

Escudero, P., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *The Journal of the Acoustical Society of America*, *128*(5), EL254-EL260. doi: 10.1121/1.3488794

Escudero, P., & Polka, L. (2003). A cross-language study of vowel categorization and vowel acoustics: Canadian English versus Canadian French. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 861–864).

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech*, *53*(3), 343–365.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalent classification. *Journal of Phonetics*, *15*, 47–65.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues* *in cross-language research* (pp. 233–277). Timonium, MD.: York Press.

Flege, J. E. (2003). Methods for assessing the perception of vowels in a second language. In *Issues in clinical linguistics* (pp. 3–28).

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*, 437–470.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*, 67–83.

Flege, J. E., Munro, M. J., & Fox, R. A. (1994). Auditory and categorical effects on cross-language vowel perception. *The Journal of the Acoustical Society of America*, *95*(6), 3623–3641.

Flege, J. E., Munro, M. J., & MacKay, I. (1995). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, *97*, 3623–31.

Flynn, S. (1987). Contrast and construction in a parameter-setting model of L2 acquisition. *Language Learning*, *37*(1), 19–62.

Fox, R. A., Flege, J. E., & Munro, M. J. (1994). The perception of English and Spanish vowels by native English and Spanish listeners: A multidimensional scaling analysis. *Journal of the Acoustic Society of America*, *97*(4), 2540–2551.

Fry, D. B., Abramson, A. S., Eimas, P. D., & Liberman, A. M. (1962). The identification and discrimination of synthetic vowels. *Language and speech*, *5*(4), 171–189.

García, P. (2014). *Perception of American English vowels by adult Spanish-English bilingual listeners* (Unpublished doctoral dissertation). Columbia University.

Gass, S. (1997). *Input, interaction, and the second language learner.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S., & Selinker, L. (2001). *Second language acquisition: an introductory course.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gerrits, E., & Schouten, M. (2004). Categorical perception depends on the discrimination task. *Perception and Psychophysics*, *66*(3), 363–376.

Giegerich, H. (1992). *English phonology: An introduction.* Cambridge: Cambridge University Press.

Gregg, K. (2001). Learnability and second language theory. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 152–180). Cambridge: Cambridge University Press.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (Ed.), *One mind, two languages: Bilingual language processing.* Oxford: Blackwell.

Hamann, S., & Colombo, I. E. (2017). A formal account of the interaction of orthography and perception. *Natural Language & Linguistic Theory*, *35*, 683–714.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*(3), 360 - 378.

Hazan, V., & Simpson, A. (2000). The effect of cue-enhancement on consonant intelligibility in noise: Speaker and listener effects. *Language and Speech*, *43*(3), 273-294.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical society of America*, *97*(5), 3099–3111.

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America*, *108*(6), 3013–3022.

Hoenigswald, H. (1960). *Language change and linguistic reconstruction.* Univ. of Chicago Press.

Holt, L., & Lotto, A. (2010). Speech perception as categorization. *Attention, perception and psychophysics*, *72*(5), 1218–1227.

Iverson, P., & Evans, B. (2007). Learning English vowels with different first-language systems: Perception of formant targets, formant movement, and duration. *Journal of the Acoustic Society of America*, *122*.

Iverson, P., & Evans, B. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *Journal of the Acoustic Society of America*, *126*(2), 866–877.

Iverson, P., & Kuhl, P. (2000). Perceptual magnet and phoneme boundary effects in speech perception: Do they arise from a common mechanism? *Perception & Psychophysics*, *62*(4), 874–886.

Iverson, P., Kuhl, P., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2002). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*, B47–B57.

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

Karpinska, M., Uchida, S., & Grenon, I. (2015). Vowel perception by listeners from different English dialects. In *18th International Congress of Phonetic Sciences, ICPhS 2015, Glasgow, UK, August 10-14, 2015.* Retrieved from https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0264.pdf

Krashen, S. (1981). *Second language acquisition and second language learning.*

Oxford: Pergamon.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, *50*(2), 93–107.

Kuhl, P. K., & Iverson, P. (1995). Chapter 4: Linguistic experience and the "perceptual magnet effect". In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121–154). Timonium, MD.: York Press.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606–608.

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

Lenneberg, E. (1967). *Biological foundations of language.* New York: Wiley.

Liberman, A., Harris, K., Hoffman, H., & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*(5), 358–368.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America*, *89*(2), 874-886.

Lord, G. (2008). Second language acquisition and first language phonological modification. In C. P. Project (Ed.), *Proceedings of the 10th hispanic linguistics symposium.* Somerville, MA.

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of english /r/ and /l/ by japanese bilinguals. *Applied Psycholinguistics*, *2*(04), 369–390.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: CUP Archive.

Major, R. (1998). Interlanguage phonology: an introduction. *Studies in Second Language Acquisition*, *20*(2), 131–137.

Martínez Celdrán, E., & Fernández Planas, A. M. (2007). *Manual de fonética española: articulaciones y sonidos de español.* Barcelona: Ariel.

McGuire, G. (2007). *Phonetic category learning* (Unpublished doctoral dissertation). Ohio State University.

Navarro Tomás, T. (1974). *Manual de pronunciación española.* Madrid: RAYCAR.

Núñez Cedeño, R., & Morales-Font, A. (1999). *Fonología generativa contemporánea de la lengua española.* Washington, DC: Georgetown Univesity Press.

Ohala, J. J. (1996). Speech perception is hearing sounds, not tongues. *The Journal of the Acoustical Society of America*, *99*(3), 1718–1725.

Pajak, B., & Levy, R. (2014). The role of abstraction in non-native speech perception. *Journal of Phonetics*, *46*, 147–160.

Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics*, *29*(2), 191 - 215. Retrieved from http://www.sciencedirect.com/science/article/pii/S0095447001901342 doi: http://dx.doi.org/10.1006/jpho.2001.0134

Pisoni, D. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, *13*(2), 253–260.

Polivanov, E. (1931). La perception des sons d'une langue étrangère. Travaux du Cercle Linguistique de Prague 4:79–96 [English translation: The subjective nature of the perceptions of language sounds]. In E. Polivanov (Ed.), *Selected works: Articles on general linguistics* (pp. 223–237). The Hague: Mouton.

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint interaction in generative grammar*. Oxford: Blackwell Publishing.

Quilis, A. (1993). *Tratado de fonología y fonética españolas* (Vol. 2). Madrid: Gredos.

Repp, B. H. (1984). Categorical perception: Issues, methods, findings. In N. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 10, pp. 243–335). London: Academic Press, Inc.

Ringe, D., & Eska, J. F. (2013). *Historical linguistics: Toward a twenty-first century reintegration*. Cambridge: Cambridge University Press.

Rosch, E. (1999). Chapter 8: Principles of categorization. In E. Margolis & S. Laurence (Eds.), *Concepts: core readings* (pp. 189–206). MIT Press.

Sadowsky, S. (2012). *Naturaleza fonética y estratificación sociolingüística de los alófonos vocálicos del castellano de Concepción (Chile)* (Unpublished doctoral dissertation). Universidad de Concepción.

Schachter, J. (1988). Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics*, *9*(3), 219–235. doi: 10.1093/applin/9.3.219

Schouten, B., Gerrits, E., & van Hessen, A. (2003). The end of categorical perception as we know it. *Speech Communication*, *41*, 71–80.

Schwartz, B., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, *12*(1), 40–72.

Selinker, L. (1972). Interlanguage. *IRAL - International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209–232.

Singleton, D. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics in Language Teaching*, *43*(4), 269–285.

Smith, M. S. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, *15*(2), 165–179.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149.

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: a working model. *Journal of Phonetics*, *39*(4), 456–466.

Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S. A., Nishi, K., & Jenkins, J. J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, *26*(4), 311–344.

Tarone, E. (1978). The phonology of interlanguage. In J. Richards (Ed.), *Understanding second and foreign language learning* (pp. 15–33). Rowly, MA: Newbury House.

Trubetzkoy, N. (1969). *Principles of phonology.* Berkeley & London: University of California Press.

van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology*, *6*, 1000.

VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Mahwah, NJ: Lawrence Erlbaum Associates.

Werker, J., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Attention, Perception, & Psychophysics*, *37*(1), 35–44.

Werker, J., & Tees, R. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustic Society of America*, *75*(6), 1866–1878.

White, L. (2003). *Second language acquisition and Universal Grammar.* Cambridge: Cambridge University Press.

Wiese, R. (2000). *The phonology of German.* Oxford: Oxford University Press.

Ying, H. (1995). What sort of input is needed for intake? *IRAL-International Review of Applied Linguistics in Language Teaching*, *33*(3), 175–194.

Young-Scholten, M. (1994). On positive evidence and ultimate attainment in L2 phonology. *Second Language Research*, *10*(3), 193–214.

APPENDIX

A

LIST OF PRAAT SCRIPTS

A.1 Vowel continuum creator (Klatt synthesis)

```
# Vowel continuum creator
# By Fernanda Barrientos
# Creates an /x-y/ 5-step vowel continuum using Klatt Grids.
# Duration: 0.11s

# Target directory

out_directory$ = "//nask.man.ac.uk/home$/Desktop/continua3"

# Defining /x/ and /y/ formant values
f1ef= 831
f2ef= 2000
pef= 210

f1ae= 1008
f2ae= 1761
pae= 196

f1as= 937
f2as= 1258
pas= 194

f1vt= 833
f2vt= 1453
pvt= 206

# Main loop - Creates 6 Klatt grids with their respective sound files.

for i from 1 to 6
Create KlattGrid... kg_asvt'i' 0 0.11 6 1 1 6 1 1 1
Add pitch point... 0.05 pas-(((pas-pvt)/5)*(i-1))
Add voicing amplitude point... 0.0 10
Add voicing amplitude point... 0.01 80
Add voicing amplitude point... 0.1 80
Add voicing amplitude point... 0.11 10
Add oral formant frequency point... 1 0.05 f1as-(((f1as-f1vt)/5)*(i-1))
Add oral formant bandwidth point... 1 0.05 30
Add oral formant frequency point... 2 0.05 f2as-(((f2as-f2vt)/5)*(i-1))
Add oral formant bandwidth point... 2 0.05 30
To Sound
Write to WAV file... 'out_directory$'/asvt_'i'.wav
endfor

for i from 1 to 6
Create KlattGrid... kg_asef'i' 0 0.11 6 1 1 6 1 1 1
Add pitch point... 0.05 pas-(((pas-pef)/5)*(i-1))
Add voicing amplitude point... 0.0 10
Add voicing amplitude point... 0.01 80
Add voicing amplitude point... 0.1 80
Add voicing amplitude point... 0.11 10
```

```
Add oral formant frequency point... 1 0.05 f1as-(((f1as-f1ef)/5)*(i-1))
Add oral formant bandwidth point... 1 0.05 30
Add oral formant frequency point... 2 0.05 f2as-(((f2as-f2ef)/5)*(i-1))
Add oral formant bandwidth point... 2 0.05 30
To Sound
Write to WAV file... 'out_directory$'/asef_'i'.wav
endfor

for i from 1 to 6
Create KlattGrid... kg_vtef'i' 0 0.11 6 1 1 6 1 1 1
Add pitch point... 0.05 pvt-(((pvt-pef)/5)*(i-1))
Add voicing amplitude point... 0.0 10
Add voicing amplitude point... 0.01 80
Add voicing amplitude point... 0.1 80
Add voicing amplitude point... 0.11 10
Add oral formant frequency point... 1 0.05 f1vt-(((f1vt-f1ef)/5)*(i-1))
Add oral formant bandwidth point... 1 0.05 30
Add oral formant frequency point... 2 0.05 f2vt-(((f2vt-f2ef)/5)*(i-1))
Add oral formant bandwidth point... 2 0.05 30
To Sound
Write to WAV file... 'out_directory$'/vtef_'i'.wav
endfor

for i from 1 to 6
Create KlattGrid... kg_asae'i' 0 0.11 6 1 1 6 1 1 1
Add pitch point... 0.05 pas-(((pas-pae)/5)*(i-1))
Add voicing amplitude point... 0.0 10
Add voicing amplitude point... 0.01 80
Add voicing amplitude point... 0.1 80
Add voicing amplitude point... 0.11 10
Add oral formant frequency point... 1 0.05 f1as-(((f1as-f1ae)/5)*(i-1))
Add oral formant bandwidth point... 1 0.05 30
Add oral formant frequency point... 2 0.05 f2as-(((f2as-f2ae)/5)*(i-1))
Add oral formant bandwidth point... 2 0.05 30
To Sound
Write to WAV file... 'out_directory$'/asae_'i'.wav
endfor

for i from 1 to 6
Create KlattGrid... kg_aevt'i' 0 0.11 6 1 1 6 1 1 1
Add pitch point... 0.05 pae-(((pae-pvt)/5)*(i-1))
Add voicing amplitude point... 0.0 10
Add voicing amplitude point... 0.01 80
Add voicing amplitude point... 0.1 80
Add voicing amplitude point... 0.11 10
Add oral formant frequency point... 1 0.05 f1ae-(((f1ae-f1vt)/5)*(i-1))
Add oral formant bandwidth point... 1 0.05 30
Add oral formant frequency point... 2 0.05 f2ae-(((f2ae-f2vt)/5)*(i-1))
Add oral formant bandwidth point... 2 0.05 30
To Sound
Write to WAV file... 'out_directory$'/aevt_'i'.wav
endfor

for i from 1 to 6
Create KlattGrid... kg_efae'i' 0 0.11 6 1 1 6 1 1 1
Add pitch point... 0.05 pef-(((pef-pae)/5)*(i-1))
Add voicing amplitude point... 0.0 10
Add voicing amplitude point... 0.01 80
Add voicing amplitude point... 0.1 80
Add voicing amplitude point... 0.11 10
Add oral formant frequency point... 1 0.05 f1ef-(((f1ef-f1ae)/5)*(i-1))
Add oral formant bandwidth point... 1 0.05 30
Add oral formant frequency point... 2 0.05 f2ef-(((f2ef-f2ae)/5)*(i-1))
Add oral formant bandwidth point... 2 0.05 30
To Sound
Write to WAV file... 'out_directory$'/efae_'i'.wav
endfor
```

A.2 Vowel continuum creator, resynthesis from previous audio file

```
#Vowel continuum maker#
#By Fernanda Barrientos + Henri Kauhanen#
#University of Manchester#

#Description: creates a 7-step continuum from a source (original vowel sound)
#and the values of a target vowel sound

#Directories

sound_directory$ = "/Users/mfbxkfb2/Desktop/definitivo"
out_directory$ = "/Users/mfbxkfb2/Desktop/continua2"

#Formant values. Enter F1 and F2 of target vowels here.

#Spanish vowels
a_f1 = 935
a_f2 = 1548
e_f1 = 609
e_f2 = 2263
i_f1 = 489
i_f2 = 2864
o_f1 = 597
o_f2 = 966
u_f1= 460
u_f2 = 982

#English vowels
as_f1 = 964
as_f2 = 1296
#ae_f1 = 844
#ae_f2 = 1060
hs_f1 = 379
hs_f2 = 1154
ic_f1 = 485
ic_f2 = 2236
vt_f1 = 814
vt_f2 = 1684

#Reading files

strings = Create Strings as file list: "list", sound_directory$ + "/*.wav"
numberOfFiles = Get number of strings


for ifile to numberOfFiles
selectObject: strings
filename$ = Get string... ifile
wav_file = Read from file: sound_directory$ + "/" + filename$

#Preliminary tasks: resample and peak scaling
wav_file_resampled = Resample... 11000 50
Scale peak... 0.99

#LPC to get source
lpc= To LPC (autocorrelation): 16, 0.025, 0.005, 50
selectObject: wav_file_resampled
plus lpc
Filter (inverse)
Rename... source

dif_f1 = 0
dif_f2 = 0
targetname$ = ""

#End 1: Calculating reference values
if filename$ == "as.wav"
dif_f1 = as_f1 - a_f1
dif_f2 = as_f2 - a_f2
targetname$ = "a"
#elsif filename$ == "ae.wav"
```

```
#dif_f1 = ae_f1 - a_f1
#dif_f2 = ae_f2 - a_f2
#targetname$ = "a"
elsif filename$ == "hs.wav"
dif_f1 = hs_f1 - o_f1
dif_f2 = hs_f2 - o_f2
targetname$ = "o"
elsif filename$ == "ic.wav"
dif_f1 = ic_f1 - i_f1
dif_f2 = ic_f2 - i_f2
targetname$ = "i"
elsif filename$ == "vt.wav"
dif_f1 = vt_f1 - a_f1
dif_f2 = vt_f2 - a_f2
targetname$ = "a"
endif

name$ = replace_regex$ (filename$, ".wav", "", 0)
for i from 1 to 7
selectObject: wav_file_resampled
formant_id = To Formant (burg): 0, 5, 5500, 0.015, 50
selectObject: formant_id

Formula (frequencies): "if row = 1 then self + (-(dif_f1/6) * (i-1)) else self fi"
Formula (frequencies): "if row = 2 then self + (-(dif_f2/6) * (i-1)) else self fi"
selectObject: "Sound source"
plus formant_id
Filter
selectObject: "Sound source_filt"
Rename... token'i'
Write to WAV file... 'out_directory$'/'name$'_'targetname$'_'i'.wav
select formant_id
Remove
endfor

endfor
select all
minus strings
Remove

for ifile to numberOfFiles
selectObject: strings
filename$ = Get string... ifile
wav_file = Read from file: sound_directory$ + "/" + filename$

#Preliminary tasks: resample and peak scaling
wav_file_resampled = Resample... 11000 50
Scale peak... 0.99

#LPC to get source
lpc= To LPC (autocorrelation): 16, 0.025, 0.005, 50
selectObject: wav_file_resampled
plus lpc
Filter (inverse)
Rename... source

dif_f1 = 0
dif_f2 = 0
targetname$ = ""

#End 2: Calculating reference values
if filename$ == "as.wav"
dif_f1 = as_f1 - o_f1
dif_f2 = as_f2 - o_f2
targetname$ = "o"
#elsif filename$ == "ae.wav"
# dif_f1 = ae_f1 - e_f1
# dif_f2 = ae_f2 - e_f2
# targetname$ = "e"
elsif filename$ == "hs.wav"
dif_f1 = hs_f1 - u_f1
dif_f2 = hs_f2 - u_f2
targetname$ = "u"
elsif filename$ == "ic.wav"
dif_f1 = ic_f1 - e_f1
dif_f2 = ic_f2 - e_f2
```

```
targetname$ = "e"
elsif filename$ == "vt.wav"
dif_f1 = vt_f1 - o_f1
dif_f2 = vt_f2 - o_f2
targetname$ = "o"
endif


name$ = replace_regex$ (filename$, ".wav", "", 0)
for i from 1 to 7
selectObject: wav_file_resampled
formant_id = To Formant (burg): 0, 5, 5500, 0.015, 50
selectObject: formant_id

Formula (frequencies): "if row = 1 then self + (-(dif_f1/6) * (i-1)) else self fi"
Formula (frequencies): "if row = 2 then self + (-(dif_f2/6) * (i-1)) else self fi"
selectObject: "Sound source"
plus formant_id
Filter
selectObject: "Sound source_filt"
Rename... token'i'
Write to WAV file... 'out_directory$'/'name$'_'targetname$'_'i'.wav
select formant_id
Remove
endfor
select all
minus strings
Remove
endfor

for ifile to numberOfFiles
selectObject: strings
filename$ = Get string... ifile
wav_file = Read from file: sound_directory$ + "/" + filename$

#Preliminary tasks: resample and peak scaling
wav_file_resampled = Resample... 11000 50
Scale peak... 0.99

#LPC to get source
lpc= To LPC (autocorrelation): 16, 0.025, 0.005, 50
selectObject: wav_file_resampled
plus lpc
Filter (inverse)
Rename... source

dif_f1 = 0
dif_f2 = 0
targetname$ = ""
#English-to-English: Calculating reference values
if filename$ == "as.wav"
dif_f1 = as_f1 - vt_f1
dif_f2 = as_f2 - vt_f2
targetname$ = "vt"
elsif filename$ == "hs.wav"
dif_f1 = hs_f1 - ic_f1
dif_f2 = hs_f2 - ic_f2
targetname$ = "ic"
endif


name$ = replace_regex$ (filename$, ".wav", "", 0)
for i from 1 to 7
selectObject: wav_file_resampled
formant_id = To Formant (burg): 0, 5, 5500, 0.015, 50
selectObject: formant_id

Formula (frequencies): "if row = 1 then self + (-(dif_f1/6) * (i-1)) else self fi"
Formula (frequencies): "if row = 2 then self + (-(dif_f2/6) * (i-1)) else self fi"
selectObject: "Sound source"
plus formant_id
Filter
selectObject: "Sound source_filt"
Rename... token'i'
Write to WAV file... 'out_directory$'/'name$'_'targetname$'_'i'.wav
select formant_id
```

```
Remove
endfor

endfor
select all
minus strings
Remove
```

A.3 Get formants from vowel files

```
#Get formants from vowel files#
#By Fernanda Barrientos#
#University of Manchester#

#Directories
sound_directory$ = "/Users/mfbxkfb2/Desktop/vowels"
out_directory$ =  "/Users/mfbxkfb2/Desktop/mystuff"

#Creates table
table_ID = Create Table with column names: "data", 0, "vowel F1 F2"

#Opening files
strings = Create Strings as file list: "list", sound_directory$ + "/*.wav"
numberOfFiles = Get number of strings

for ifile to numberOfFiles
selectObject: strings
fileName$ = Get string: ifile
wav_file_ID = Read from file: sound_directory$ + "/" + fileName$

#Creates formant objects to get values
formant_ID = To Formant (burg): 0.0, 5, 5500, 0.025, 50

selectObject: formant_ID
f1 = Get value at time: 1, 0.05, "Hertz", "Linear"
f2 = Get value at time: 2, 0.05, "Hertz", "Linear"

#Filling out the table with corresponding values
selectObject: table_ID
Append row
Set string value: ifile, "vowel", fileName$
Set numeric value: ifile, "F1", f1
Set numeric value: ifile, "F2", f2
endfor

#Saving table as file (will appear on
selectObject: table_ID
Save as comma-separated file: 'out_directory$'/voweldata.csv
```

A.4 Concatenator

```
#Concatenator#
#By Fernanda Barrientos - University of Manchester#
#Description: concatenates CVC noncewords + silence  + target vowel, according to
# vowel files#

# Directories: the sound directory has only vowels. Onsets and codas are specified for
#each vowel as individual paths.

sound_directory$ = "/Volumes/KINGSTON/vowels"
out_directory$ = "/Users/mfbxkfb2/Desktop/files"

# String list
Create Strings as file list... list 'sound_directory$'/*.wav
noFiles = Get number of strings

# Main loop
for ifile to noFiles
select Strings list
filename$ = Get string... ifile

vowel$ = left$ (filename$, 7)

#if loop: defines which onsets and codas go with which vowels.

if vowel$ == "pot1_a_"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/book1_b_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/cup2_p_1.wav

elsif vowel$ == "pot1_o_"
# Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/book1_b_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/cup2_p_1.wav

elsif vowel$ == "pit1_i_"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/peak2_p_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/cup2_p_1.wav

elsif vowel$ == "pit1_e_"
 #Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/peak2_p_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/cup2_p_1.wav

elsif vowel$ == "book1_o"
 #Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/pot1_p_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/peak2_k_1.wav

elsif vowel$ == "book1_u"
# Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/pot1_p_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/peak2_k_1.wav
```

```
elsif vowel$ == "cut1_a_"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/pot1_p_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/beat2_t_1.wav

elsif vowel$ == "cut1_o_"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/pot1_p_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/peak2_k_1.wav

elsif vowel$ == "pot1_vt"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/book1_b_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/cup2_p_1.wav

elsif vowel$ == "book1_i"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/pot1_p_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/cup2_p_1.wav

elsif vowel$ == "book1_a"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/book1_b_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/beat2_t_1.wav

elsif vowel$ == "pit1_as"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/beat2_b_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/peak2_k_1.wav

elsif vowel$ == "pit1_vt"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/beat2_b_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/peak2_k_1.wav

elsif vowel$ == "cut1_hs"
#Onset
Read from file... /Volumes/KINGSTON/vowels/onsets/beat2_b_1.wav
# Vowel
Read from file... 'sound_directory$'/'filename$'
# Coda
Read from file... /Volumes/KINGSTON/vowels/codas/cup2_p_1.wav

endif


# Creates Silence
#Create Sound from formula... silence 1 0 1 11000 0

# Choosing target vowel
#vowel$ = left$ (filename$, 4)
#if vowel$ == "ae_a"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/ae_a_1.wav
#elsif vowel$ == "ae_e"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/ae_e_1.wav
```

```
#elsif vowel$ == "as_a"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/as_a_1.wav
#elsif vowel$ == "as_o"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/as_o_1.wav
#elsif vowel$ == "ic_i"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/ic_i_1.wav
#elsif vowel$ == "ic_e"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/ic_e_1.wav
#elsif vowel$ == "hs_o"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/hs_o_1.wav
#elsif vowel$ == "hs_u"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/hs_u_1.wav
#elsif vowel$ == "vt_a"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/vt_a_1.wav
#elsif vowel$ == "vt_o"
#Read from file... /Users/mfbxkfb2/Desktop/continua3/desilenced/vt_o_1.wav
#endif

# Concatenate everything
select all
minus Strings list
Concatenate

# Write
Write to WAV file... 'out_directory$'/'vowel$''ifile'.wav

# Remove before next iteration of for loop
select all
minus Strings list
Remove

endfor

# Tidy up
select all
Remove
echo all CVC files concatenated.
```

APPENDIX

B

LANGUAGE PROFICIENCY QUESTIONNAIRES

B.1 Experiment 1

English proficiency questionnaire

Subject Number: ______

Part I. Please circle the number that you consider more appropriate. (1= I can't do it; 5= very well).

1. How would you rate your ability to understand **spoken** British English?

   1 2 3 4 5

2. How would you rate your ability to understand **spoken** American English?

   1 2 3 4 5

3. How would you rate your ability to **read** English texts?

   1 2 3 4 5

Part II. Please circle all that apply.

1. How did you learn English?
   a. I attended an immersion school when I was young
   b. I took English classes
   c. Self- taught
   d. It was spoken in my house on a regular basis
   e. Other:

2. When did you start learning English?
   a. Before I was 5 years old
   b. Between 5 and 12 years old
   c. Between 12 and 18 years old
   d. After I was 18

3. Did you ever travel to an English-speaking country for over a month before coming to Manchester?
   a. Yes –How old were you (if several times, just think of your first time) _________
   b. No

4. Do you speak any other language fluently?
   a. Yes –Which one (s)? ____
   b. No

5. If you have any additional comments regarding these questions please write them down here:
   ____
   ____
   ____
   ____

   Thanks!! ☺

B.2 Experiment 2

Language questionnaire

Please read the Participant Information Sheet first (it will open in a new tab). If you are happy to participate please tick all the following boxes:

☐ I confirm that I am over 18 years of age.
☐ I confirm that I have read the information sheet and have had the opportunity to reconsider the information and ask questions, and had these answered satisfactorily.
☐ I agree that any data collected may be used for the purposes of research and that the information collected may be passed as anonymous data to other researchers.

Before we begin, I would like to ask you a few questions.
* What is your native language?
Spanish
English

Have you taken a standardized English proficiency test such as TOEFL, IELTS, etc? If you did, please write down your overall score, and the name of the test if you ticked "other". If you took the same test more than once, or took two different tests, then choose the latest.
◯ TOEFL
◯ IELTS
◯ Other (TOEIC, BULATS, etc).
◯ None

Please enter your score here: ___________

Have you ever lived in an English speaking country for over a month? Enter the amount of time in months. If you have not done so, then please enter 0. _________

On a scale of 1 to 5, how would you rate your SPOKEN proficiency in English? 1=I can't speak/understand anything at all, and 5=I can speak/understand it very fluently.

◯ 1
◯ 2
◯ 3
◯ 4
◯ 5

APPENDIX

C

TABLES WITH STATISTICAL VALUES

C.1 Chapter 4: d-prime *(d')*

Table C.1: Sensitivity (d') and standard errors (SE).

	/ɑ-ʌ/		/ɑ-a/		/ɑ-o/		/ʌ-a/		/ʌ-o/	
	d'	SE	d'	SE	d'	SE	d'	SE	d'	SE
NNS-B	2.57	0.67	0.88	0.88	6.28	1.09	0.87	0.88	4.70	0.75
NNS-A	3.93	0.88	1.87	0.91	5.14	0.97	0.00	0.00	5.23	0.99
NS	4.99	0.99	2.43	0.90	5.75	1.16	2.43	0.90	5.75	1.16

C.2 Chapter 4: Logistic regressions

C.2.1 Section 4.5.2.1

C.2.1.1 /ɑ - o/ continuum

Table C.2: Analysis of deviance. Stimuli from /ɑ - o/ continuum (stim1) discriminated against /ɑ/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			208	288.93		
stim1	6	141.36	202	147.56	<2e-16	***
Level	2	1.22	200	146.35	0.5446	
stim1:Level	12	9.41	188	136.94	0.6674	

Table C.3: GLM results, stimuli from /ɑ - o/ continuum (stim1) discriminated against /ɑ/, with stimulus as a single predictor.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	0.2683	0.3684	0.73	0.4665	
stim11	-18.8343	1190.8655	-0.02	0.9874	
stim12	-2.9073	0.8194	-3.55	0.0004	***
stim13	-2.1401	0.6513	-3.29	0.0010	**
stim15	0.5302	0.5448	0.97	0.3305	
stim16	1.6035	0.6513	2.46	0.0138	*
stim17	3.0990	1.0818	2.86	0.0042	**

Table C.4: /ɑ-o/ - /ɑ/ discrimination task, prediction: Table of observed values, fitted values, and their corresponding standard error.

stim1	level	observed	fitted	se
1	NNSA	0.0000	0.0000	0.0000
1	NNSB	0.0000	0.0000	0.0000
1	NS	0.0000	0.0000	0.0000
2	NNSA	0.2222	0.0667	0.0455
2	NNSB	0.0000	0.0667	0.0455
2	NS	0.0000	0.0667	0.0455
3	NNSA	0.1111	0.1333	0.0621
3	NNSB	0.1538	0.1333	0.0621
3	NS	0.1250	0.1333	0.0621
4	NNSA	0.6667	0.5667	0.0905
4	NNSB	0.5385	0.5667	0.0905
4	NS	0.5000	0.5667	0.0905
5	NNSA	0.7778	0.6897	0.0859
5	NNSB	0.5833	0.6897	0.0859
5	NS	0.7500	0.6897	0.0859
6	NNSA	0.8889	0.8667	0.0621
6	NNSB	0.9231	0.8667	0.0621
6	NS	0.7500	0.8667	0.0621
7	NNSA	0.8889	0.9667	0.0328
7	NNSB	1.0000	0.9667	0.0328
7	NS	1.0000	0.9667	0.0328

Table C.5: Analysis of deviance. Stimuli from /ɑ - o/ continuum (stim1) discriminated against /o/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			209	290.65		
Level	2	2.17	207	288.47	0.3372	
stim1	6	180.67	201	107.81	<2e-16	***
Level:stim1	12	9.54	189	98.26	0.6561	

Table C.6: GLM results, stimuli from /ɑ - o/ continuum (stim1) discriminated against /o/ with stimulus as a single predictor.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.1335	0.3660	-0.36	0.7152	
stim11	19.6996	1963.4052	0.01	0.9920	
stim12	2.7726	0.8183	3.39	0.0007	***
stim13	1.5198	0.5850	2.60	0.0094	**
stim15	-2.0637	0.7101	-2.91	0.0037	**
stim16	-3.2338	1.0809	-2.99	0.0028	**
stim17	-19.4325	1963.4052	-0.01	0.9921	

Table C.7: /ɑ-o/ - /o/ discrimination task, prediction: Table of observed values, fitted values, and their corresponding standard error.

stim1	level	observed	fitted	se
1	NNSA	1.0000	1.0000	0.0000
1	NNSB	1.0000	1.0000	0.0000
1	NS	1.0000	1.0000	0.0000
2	NNSA	0.8889	0.9333	0.0455
2	NNSB	0.9231	0.9333	0.0455
2	NS	1.0000	0.9333	0.0455
3	NNSA	0.8889	0.8000	0.0730
3	NNSB	0.7692	0.8000	0.0730
3	NS	0.7500	0.8000	0.0730
4	NNSA	0.5556	0.4667	0.0911
4	NNSB	0.3077	0.4667	0.0911
4	NS	0.6250	0.4667	0.0911
5	NNSA	0.0000	0.1000	0.0548
5	NNSB	0.0000	0.1000	0.0548
5	NS	0.3750	0.1000	0.0548
6	NNSA	0.0000	0.0333	0.0328
6	NNSB	0.0000	0.0333	0.0328
6	NS	0.1250	0.0333	0.0328
7	NNSA	0.0000	0.0000	0.0000
7	NNSB	0.0000	0.0000	0.0000
7	NS	0.0000	0.0000	0.0000

C.2.1.2 /ɑ - a/ continuum

Table C.8: Analysis of deviance. Stimuli from /ɑ - a/ continuum (stim1) discriminated against /ɑ/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			206	80.11		
stim1	6	16.41	200	63.70	0.0117	*
Level	2	3.90	198	59.80	0.1421	
stim1:Level	12	5.78	186	54.02	0.9266	

Table C.9: GLM results, stimuli from /ɑ - a/ continuum (stim1) discriminated against /ɑ/, with stimulus as a single predictor.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-3.3673	1.0171	-3.31	0.0009	***
stim11	-17.1988	3292.4472	-0.01	0.9958	
stim12	-17.1988	3237.1080	-0.01	0.9958	
stim13	0.0000	1.4384	0.00	1.0000	
stim15	-17.1988	3292.4472	-0.01	0.9958	
stim16	1.1701	1.1853	0.99	0.3236	
stim17	1.7987	1.1297	1.59	0.1113	

Table C.10: /ɑ - a/ - /ɑ/ discrimination task: Observed and vitted values with standard errors.

stim1	level	observed	fitted	se
1	NNSA	0.0000	0.0000	0.0000
1	NNSB	0.0000	0.0000	0.0000
1	NS	0.0000	0.0000	0.0000
2	NNSA	0.0000	0.0000	0.0000
2	NNSB	0.0000	0.0000	0.0000
2	NS	0.0000	0.0000	0.0000
3	NNSA	0.0000	0.0333	0.0328
3	NNSB	0.0000	0.0333	0.0328
3	NS	0.1250	0.0333	0.0328
4	NNSA	0.0000	0.0333	0.0328
4	NNSB	0.0769	0.0333	0.0328
4	NS	0.0000	0.0333	0.0328
5	NNSA	0.0000	0.0000	0.0000
5	NNSB	0.0000	0.0000	0.0000
5	NS	0.0000	0.0000	0.0000
6	NNSA	0.0000	0.1000	0.0548
6	NNSB	0.1538	0.1000	0.0548
6	NS	0.1250	0.1000	0.0548
7	NNSA	0.1111	0.1724	0.0701
7	NNSB	0.0833	0.1724	0.0701
7	NS	0.3750	0.1724	0.0701

Table C.11: Analysis of deviance. Stimuli from /ɑ - a/ continuum (stim1) discriminated against /a/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			209	122.85		
stim1	6	11.12	203	111.73	0.0847	.
Level	2	4.68	201	107.05	0.0961	.
stim1:Level	12	6.53	189	100.52	0.8873	

Table C.12: GLM results, stimuli from /ɑ - a/ continuum (stim1) discriminated against /a/, with level as a single predictor.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-3.0796	0.5113	-6.02	1.71e-09	***
Level2	0.8283	0.6676	1.24	0.2147	
Level3	1.2879	0.6382	2.02	0.0436	*

Table C.13: /ɑ - a/ - /a/ discrimination task: predicted and fitted values with standard errors.

stim1	level	observed	fitted	se
1	NNSA	0.2222	0.0952	0.0370
1	NNSB	0.0769	0.0440	0.0215
1	NS	0.2500	0.1429	0.0468
2	NNSA	0.1111	0.0952	0.0370
2	NNSB	0.0769	0.0440	0.0215
2	NS	0.3750	0.1429	0.0468
3	NNSA	0.1111	0.0952	0.0370
3	NNSB	0.0769	0.0440	0.0215
3	NS	0.1250	0.1429	0.0468
4	NNSA	0.1111	0.0952	0.0370
4	NNSB	0.0000	0.0440	0.0215
4	NS	0.1250	0.1429	0.0468
5	NNSA	0.1111	0.0952	0.0370
5	NNSB	0.0000	0.0440	0.0215
5	NS	0.1250	0.1429	0.0468
6	NNSA	0.0000	0.0952	0.0370
6	NNSB	0.0000	0.0440	0.0215
6	NS	0.0000	0.1429	0.0468
7	NNSA	0.0000	0.0952	0.0370
7	NNSB	0.0769	0.0440	0.0215
7	NS	0.0000	0.1429	0.0468

C.2.2 Section 4.5.2.2

C.2.2.1 /ʌ - o/ continuum

Table C.14: Analysis of deviance. Stimuli from /ʌ - o/ continuum (stim1) discriminated against /ʌ/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			206	271.06		
stim1	6	142.84	200	128.23	< 2e-16	***
Level	2	7.66	198	120.56	0.0217	*
stim1:Level	12	11.69	186	108.87	0.4706	

Table C.15: GLM results, stimuli from /ʌ - o/ continuum (stim1) discriminated against /ʌ/, with level and stimulus as predictors.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	2.4242	0.7063	3.43	0.0006	***
stim11	-4.3666	0.8676	-5.03	4.84e-07	***
stim12	-4.4038	0.8658	-5.09	3.65e-07	***
stim13	-2.8994	0.7611	-3.81	0.0001	***
stim15	0.4581	0.9689	0.47	0.6364	
stim16	1.1749	1.2003	0.98	0.3277	
stim17	16.3503	1151.5717	0.01	0.9887	
Level2	0.7552	0.5590	1.35	0.1767	
Level3	-1.0246	0.6170	-1.66	0.0968	.

Table C.16: /ʌ-o/ - /ʌ/ discrimination task, prediction: Table of observed values, fitted values, and their corresponding standard error.

stim1	level	observed	fitted	se
1	NNSA	0.1111	0.2338	0.1099
1	NNSB	0.2500	0.1254	0.0689
1	NS	0.0000	0.0489	0.0354
2	NNSA	0.2222	0.2272	0.1080
2	NNSB	0.0769	0.1214	0.0664
2	NS	0.1250	0.0472	0.0342
3	NNSA	0.7500	0.5695	0.1323
3	NNSB	0.3077	0.3834	0.1107
3	NS	0.1250	0.1825	0.0922
4	NNSA	1.0000	0.9601	0.0301
4	NNSB	0.9231	0.9187	0.0528
4	NS	0.7500	0.8021	0.1094
5	NNSA	0.8889	0.9744	0.0222
5	NNSB	0.9231	0.9470	0.0412
5	NS	1.0000	0.8650	0.0928
6	NNSA	1.0000	0.9873	0.0143
6	NNSB	1.0000	0.9734	0.0283
6	NS	0.8750	0.9292	0.0696
7	NNSA	1.0000	1.0000	0.0000
7	NNSB	1.0000	1.0000	0.0000
7	NS	1.0000	1.0000	0.0000

Table C.17: Analysis of deviance. Stimuli from /ʌ - o/ continuum (stim1) discriminated against /o/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			208	288.66		
stim1	6	130.26	202	158.40	< 2e-16	***
Level	2	9.13	200	149.27	0.0104	*
stim1:Level	12	9.57	188	139.70	0.6538	

Table C.18: GLM results, stimuli from /ʌ - o/ continuum (stim1) discriminated against /o/, with level and stimulus as predictors.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.7350	0.4558	-1.61	0.1068	
stim11	3.6641	1.0961	3.34	0.0008	***
stim12	2.9270	0.8377	3.49	0.0005	***
stim13	1.6361	0.6094	2.68	0.0073	**
stim15	-0.1031	0.5482	-0.19	0.8508	
stim16	-1.8824	0.6797	-2.77	0.0056	**
stim17	-18.5700	1151.2000	-0.02	0.9871	
Level2	0.6264	0.4823	1.30	0.1941	
Level3	1.5416	0.5360	2.88	0.0040	**

Table C.19: /ʌ-o/ - /o/ discrimination task, prediction: Table of observed values, fitted values, and their corresponding standard error.

stim1	level	observed	fitted	se
1	NNSA	1.0000	0.9722	0.0291
1	NNSB	0.9231	0.9493	0.0497
1	NS	1.0000	0.9887	0.0125
2	NNSA	1.0000	0.9437	0.0432
2	NNSB	0.8462	0.8995	0.0682
2	NS	1.0000	0.9767	0.0197
3	NNSA	0.7778	0.8216	0.0832
3	NNSB	0.6923	0.7112	0.1027
3	NS	1.0000	0.9200	0.0468
4	NNSA	0.3333	0.4729	0.1208
4	NNSB	0.3077	0.3241	0.0998
4	NS	0.8750	0.6914	0.1134
5	NNSA	0.4444	0.4473	0.1207
5	NNSB	0.4167	0.3019	0.0990
5	NS	0.5000	0.6690	0.1175
6	NNSA	0.2222	0.1201	0.0681
6	NNSB	0.0769	0.0680	0.0407
6	NS	0.1250	0.2543	0.1160
7	NNSA	0.0000	0.0000	0.0000
7	NNSB	0.0000	0.0000	0.0000
7	NS	0.0000	0.0000	0.0000

C.2.2.2 /ʌ - a/ continuum

Table C.20: Analysis of deviance. Stimuli from /ʌ - a/ continuum (stim1) discriminated against /ʌ/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			208	117.89	
stim1	6	8.51	202	109.38	0.2032
Level	2	3.73	200	105.65	0.1549
stim1:Level	12	12.18	188	93.47	0.4312

Table C.21: GLM results, stimuli from /ʌ - a/ continuum (stim1) discriminated against /ʌ/, with level as single predictor.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-2.4849	0.3934	-6.32	0.0000
Level2	0.5754	0.5462	1.05	0.2921
Level3	-0.8109	0.8205	-0.99	0.3230

Table C.22: /ʌ-a/ - /ʌ/ discrimination task. Observed and fitted values with standard errors.

stim1	level	observed	fitted	se
1	NNSA	0.2222	0.1290	0.0426
1	NNSB	0.0000	0.0769	0.0279
1	NS	0.0000	0.0357	0.0248
2	NNSA	0.0000	0.1290	0.0426
2	NNSB	0.0000	0.0769	0.0279
2	NS	0.0000	0.0357	0.0248
3	NNSA	0.1111	0.1290	0.0426
3	NNSB	0.0000	0.0769	0.0279
3	NS	0.0000	0.0357	0.0248
4	NNSA	0.1111	0.1290	0.0426
4	NNSB	0.1538	0.0769	0.0279
4	NS	0.0000	0.0357	0.0248
5	NNSA	0.1250	0.1290	0.0426
5	NNSB	0.1538	0.0769	0.0279
5	NS	0.0000	0.0357	0.0248
6	NNSA	0.2222	0.1290	0.0426
6	NNSB	0.1538	0.0769	0.0279
6	NS	0.0000	0.0357	0.0248
7	NNSA	0.1111	0.1290	0.0426
7	NNSB	0.0769	0.0769	0.0279
7	NS	0.2500	0.0357	0.0248

Table C.23: Analysis of deviance. Stimuli from /ʌ - a/ continuum (stim1) discriminated against /a/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			209	80.41		
stim1	6	12.37	203	68.04	0.0542	.
Level	2	0.07	201	67.96	0.9635	
stim1:Level	12	10.34	189	57.63	0.5866	

Table C.24: GLM results, stimuli from /ʌ - a/ continuum (stim1) discriminated against /a/, with stimulus as single predictor.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-19.5661	1963.4052	-0.01	0.9920
stim11	17.9566	1963.4052	0.01	0.9927
stim12	16.9270	1963.4053	0.01	0.9931
stim13	16.1988	1963.4054	0.01	0.9934
stim15	16.1988	1963.4054	0.01	0.9934
stim16	0.0000	2776.6742	0.00	1.0000
stim17	16.1988	1963.4054	0.01	0.9934

Table C.25: /ʌ-a/ - /a/ discrimination task. Observed and fitted values with standard errors.

stim1	level	observed	fitted	se
1	NNSA	0.1111	0.1667	0.0680
1	NNSB	0.0769	0.1667	0.0680
1	NS	0.3750	0.1667	0.0680
2	NNSA	0.1111	0.0667	0.0455
2	NNSB	0.0769	0.0667	0.0455
2	NS	0.0000	0.0667	0.0455
3	NNSA	0.0000	0.0333	0.0328
3	NNSB	0.0769	0.0333	0.0328
3	NS	0.0000	0.0333	0.0328
4	NNSA	0.0000	0.0000	0.0000
4	NNSB	0.0000	0.0000	0.0000
4	NS	0.0000	0.0000	0.0000
5	NNSA	0.1111	0.0333	0.0328
5	NNSB	0.0000	0.0333	0.0328
5	NS	0.0000	0.0333	0.0328
6	NNSA	0.0000	0.0000	0.0000
6	NNSB	0.0000	0.0000	0.0000
6	NS	0.0000	0.0000	0.0000
7	NNSA	0.0000	0.0333	0.0328
7	NNSB	0.0769	0.0333	0.0328
7	NS	0.0000	0.0333	0.0328

C.2.3 Section 4.5.2.3

Table C.26: Analysis of deviance. Stimuli from /ɑ - ʌ/ continuum (stim1) discriminated against /ɑ/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			206	226.57		
stim1	6	80.05	200	146.52	3.487e-15	***
Level	2	30.82	198	115.69	2.026e-07	***
stim1:Level	12	8.38	186	107.31	0.7544	

Table C.27: GLM results, stimuli from /ɑ - ʌ/ continuum (stim1) discriminated against /ɑ/, with level and stimulus as predictors.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-4.6734	0.9685	-4.83	1.40e-06	***
stim11	-0.7867	1.2991	-0.61	0.5448	
stim12	-15.8517	1085.9527	-0.01	0.9884	
stim13	-0.7867	1.2991	-0.61	0.5448	
stim15	2.5003	0.9410	2.66	0.0079	**
stim16	3.5440	0.9676	3.66	0.0002	***
stim17	4.5096	1.0005	4.51	6.56e-06	***
Level2	1.3123	0.5684	2.31	0.0209	*
Level3	3.2697	0.7046	4.64	3.47e-06	***

Table C.28: /ɑ-ʌ/ - /ɑ/ discrimination task, prediction: Table of observed values, fitted values, and their corresponding standard error.

stim1	Level	Observed prop.	Fitted prob.	SE
1	NNSA	0.0000	0.0156	0.0180
1	NNSB	0.0769	0.0042	0.0051
1	NS	0.0000	0.1006	0.0953
2	NNSA	0.0000	0.0000	0.0000
2	NNSB	0.0000	0.0000	0.0000
2	NS	0.0000	0.0000	0.0000
3	NNSA	0.0000	0.0156	0.0180
3	NNSB	0.0000	0.0042	0.0051
3	NS	0.1250	0.1006	0.0953
4	NNSA	0.0000	0.0335	0.0302
4	NNSB	0.0000	0.0093	0.0089
4	NS	0.2500	0.1972	0.1246
5	NNSA	0.3333	0.2972	0.1167
5	NNSB	0.0769	0.1022	0.0563
5	NS	0.7500	0.7496	0.1178
6	NNSA	0.5556	0.5456	0.1305
6	NNSB	0.2500	0.2443	0.0986
6	NS	0.8750	0.8948	0.0651
7	NNSA	0.7778	0.7592	0.1053
7	NNSB	0.4167	0.4591	0.1236
7	NS	1.0000	0.9571	0.0306

Table C.29: Analysis of deviance. Stimuli from /ɑ - ʌ/ continuum (stim1) discriminated against /ʌ/.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			208	240.91		
stim1	6	62.76	202	178.15	1.234e-11	***
Level	2	21.56	200	156.58	2.076e-05	***
stim1:Level	12	7.92	188	148.66	0.7912	

Table C.30: GLM results, stimuli from /ɑ - ʌ/ continuum (stim1) discriminated against /ʌ/, with level and stimulus as predictors.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-2.0192	0.5467	-3.69	0.0002	***
stim11	1.6946	0.6170	2.75	0.0060	**
stim12	1.3756	0.6109	2.25	0.0243	*
stim13	0.3819	0.6203	0.62	0.5382	
stim15	-1.7285	0.8841	-1.96	0.0506	.
stim16	-2.5859	1.1325	-2.28	0.0224	*
stim17	-17.6722	1108.3051	-0.02	0.9873	
Level2	0.8216	0.4735	1.74	0.0827	.
Level3	2.2146	0.5157	4.29	1.76e-05	***

Table C.31: Prediction: /ɑ-ʌ/ - /ʌ/ discrimination task.

stim1	level	observed	fitted	se
1	NNSA	0.6667	0.6218	0.1158
1	NNSB	0.3846	0.4196	0.1102
1	NS	0.8750	0.8688	0.0641
2	NNSA	0.5556	0.5444	0.1201
2	NNSB	0.3077	0.3444	0.1034
2	NS	0.8750	0.8279	0.0781
3	NNSA	0.2222	0.3067	0.1079
3	NNSB	0.1538	0.1628	0.0699
3	NS	0.7500	0.6404	0.1206
4	NNSA	0.1111	0.2319	0.0954
4	NNSB	0.2308	0.1172	0.0566
4	NS	0.5000	0.5487	0.1306
5	NNSA	0.1111	0.0509	0.0401
5	NNSB	0.0000	0.0230	0.0189
5	NS	0.1429	0.1775	0.1139
6	NNSA	0.1111	0.0222	0.0239
6	NNSB	0.0000	0.0099	0.0108
6	NS	0.0000	0.0839	0.0803
7	NNSA	0.0000	0.0000	0.0000
7	NNSB	0.0000	0.0000	0.0000
7	NS	0.0000	0.0000	0.0000

C.3 Chapter 5: t-tests for rating tasks

Table C.32: Comparing mean ratings of stimuli categorised as /ɑ/ vs. /ʌ/, L2-label task. (t-test: unpaired, two-tailed)

	t-test	Df	p-value	
NNS-B	-5.809	230.93	2.069e-08	***
NNS-A	-0.883	122.75	0.3789	
NS	3.543	235.85	0.0005	***

Table C.33: Comparing mean ratings of /ɛ/ vs. /ʌ/, L2-label task. (t-test: unpaired, two-tailed)

	t-test	Df	p-value	
NNS-B	-3.989	268.38	8.531e-05	***
NNS-A	-2.369	122.95	0.0194	*
NS	-2.584	241.69	0.0104	*

Table C.34: Comparing mean ratings of stimuli categorised as /a/ vs. /o/, L1-label task. (t-test: unpaired, two-tailed)[1]

	t-test	Df	p-value	
NNS-B	-4.812	43.61	1.819e-05	***
NNS-A	-5.623	41.83	1.4e-06	***

Table C.35: Comparing mean ratings of stimuli categorised as /a/ vs. /e/, L1-label task. (t-test: unpaired, two-tailed)

	t-test	Df	p-value
NNS-B	-1.455	219.71	0.1472
NNS-A	-1.0721	144.29	0.2855

[1]These t-tests were not taken into account due to the large differences in sample size.

Table C.36: Across-group comparisons (NNS-B vs. NNS-A), L1-label task, /ɑ-ʌ/ continuum. (t-test: unpaired, two-tailed)

	t-test	Df	p-value	
Stimuli perceived as /a/	-4.887	436.35	1.438e-06	***
Stimuli perceived as /o/	-2.175	58.11	0.0337	*

Table C.37: Across-group comparisons (NNS-B vs. NNS-A), L1-label task, /ʌ-ɛ/ continuum. (t-test: unpaired, two-tailed)

	t-test	Df	p-value	
Stimuli perceived as /a/	-2.421	155.92	0.0166	*
Stimuli perceived as /e/	-2.127	204.81	0.0346	*

C.4 Chapter 5: d-prime (*d'*)

Table C.38: Mean d-prime scores for each stimuli pair, /ɑ - ʌ/ continuum, 1-step discrimination task.

Group	Stim. pair	Mean d'	St. Error
NNS-B	1-2	0.3686	0.1903
NNS-B	2-3	0.8525	0.418
NNS-B	3-4	1.2823	0.3918
NNS-B	4-5	1.4750	0.4395
NNS-A	1-2	0.2914	0.2914
NNS-A	2-3	1.2613	0.358
NNS-A	3-4	0.9648	0.485
NNS-A	4-5	1.3379	0.4605
NS	1-2	1.0384	0.3433
NS	2-3	1.3234	0.5122
NS	3-4	1.5035	0.5127
NS	4-5	0.9557	0.3577

Table C.39: Kruskal-Wallis test, for *d'* scores, /ɑ - ʌ/ continuum (1-step).

Stim. pair	Chi-square	df	p-value
1-2	4.6218	2	0.0992
2-3	1.2760	2	0.5284
3-4	0.8610	2	0.6502
4-5	0.4938	2	0.7812

Table C.40: Mean d-prime scores for each stimuli pair, /ɑ - ʌ/ continuum, 2-step discrimination task.

Group	Stim. pair	Mean d'	St. Error
NNS-B	1-3	2.2429	0.2228
NNS-B	2-4	2.8031	0.3118
NNS-B	3-5	2.4382	0.4268
NNS-A	1-3	2.0249	0.5895
NNS-A	2-4	2.6011	0.4954
NNS-A	3-5	2.7512	0.4127
NS	1-3	2.7785	0.4914
NS	2-4	3.1273	0.4051
NS	3-5	3.1896	0.486

Table C.41: Kruskal-Wallis test, for *d'* scores, /ɑ - ʌ/ continuum (2-step).

Stim. pair	Chi-square	df	p-value
1-3	1.1932	2	0.5507
2-4	0.9640	2	0.6176
3-5	1.8366	2	0.3992

Table C.42: Mean d-prime scores for each stimuli pair, /ʌ - ɛ/ continuum, 1-step discrimination task.

Group	Stim. pair	Mean d'	St. Error
NNS-B	1-2	1.3856	0.3747
NNS-B	2-3	1.428	0.4808
NNS-B	3-4	0.9367	0.3688
NNS-B	4-5	0.9067	0.3367
NNS-A	1-2	1.9377	0.5483
NNS-A	2-3	2.3818	0.3822
NNS-A	3-4	1.1603	0.4319
NNS-A	4-5	1.3432	0.4988
NS	1-2	2.0874	0.4938
NS	2-3	2.0933	0.5033
NS	3-4	2.2029	0.6225
NS	4-5	2.1922	0.4171

Table C.43: Kruskal-Wallis test, for *d'* scores, /ʌ - ɛ/ continuum (1-step).

Stim. pair	Chi-square	df	p-value
1-2	1.7480	2	0.4173
2-3	2.0224	2	0.3638
3-4	2.6538	2	0.2653
4-5	3.9591	2	0.1381

Table C.44: Mean d-prime scores for each stimuli pair, /ʌ - ɛ/ continuum, 2-step discrimination task.

Group	Stim. pair	Mean d'	St. Error
NNS-B	1-3	2.4127	0.4619
NNS-B	2-4	2.5743	0.5261
NNS-B	3-5	1.8925	0.5261
NNS-A	1-3	3.5031	0.462
NNS-A	2-4	3.4301	0.4721
NNS-A	3-5	3.5283	0.2761
NS	1-3	3.1066	0.4508
NS	2-4	3.7327	0.3672
NS	3-5	3.5949	0.4196

Table C.45: Kruskal-Wallis test, for *d'* scores, /ʌ - ɛ/ continuum (2-step).

Stim. pair	Chi-square	df	p-value
1-3	3.3880	2	0.1838
2-4	5.0400	2	0.0805
3-5	8.0920	2	0.0175

C.5 Chapter 5: Logistic regressions

C.5.1 L2-label categorisation task

C.5.1.1 /ɑ-ʌ continuum/

Table C.46: Analysis of deviance, L2-label categorisation task: /ɑ-ʌ/ continuum.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			749	1009.52		
Group	2	33.51	747	976.01	5.285e-08	***
stim	4	38.83	743	937.17	7.541e-08	***
Group:stim	8	29.21	735	907.97	0.0003	***

Table C.47: GLM results, L2-label categorisation task: /ɑ-ʌ/ continuum. Interaction between group and stimulus as predictor.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-0.7949	0.3220	-2.47	0.0136	*
GroupNNSB	0.0744	0.4316	0.17	0.8632	
GroupNS	1.9476	0.4619	4.22	2.48e-05	***
stimasvt2	0.4813	0.4413	1.09	0.2755	
stimasvt3	-0.3335	0.4733	-0.70	0.4810	
stimasvt4	-0.3335	0.4733	-0.70	0.4810	
stimasvt5	-0.4578	0.4819	-0.95	0.3421	
GroupNNSB:stimasvt2	-0.0167	0.5927	-0.03	0.9776	
GroupNS:stimasvt2	-0.2477	0.6553	-0.38	0.7055	
GroupNNSB:stimasvt3	0.4945	0.6206	0.80	0.4256	
GroupNS:stimasvt3	-0.4137	0.6457	-0.64	0.5218	
GroupNNSB:stimasvt4	0.5722	0.6194	0.92	0.3555	
GroupNS:stimasvt4	-1.3945	0.6484	-2.15	0.0315	*
GroupNNSB:stimasvt5	0.3740	0.6324	0.59	0.5543	
GroupNS:stimasvt5	-2.2112	0.6909	-3.20	0.0014	**

Table C.48: GLM Prediction: /ɑ/ (*pot*) responses for each /ɑ - ʌ/ continuum step. Observed proportions for each response, fitted values, and standard error.

Group	stim	pot	cup	fitted	se
NNSB	1	0.33	0.67	0.33	0.06
NNSB	2	0.44	0.56	0.44	0.07
NNSB	3	0.36	0.64	0.36	0.06
NNSB	4	0.38	0.62	0.38	0.07
NNSB	5	0.31	0.69	0.31	0.06
NNSA	1	0.31	0.69	0.31	0.07
NNSA	2	0.42	0.58	0.42	0.07
NNSA	3	0.24	0.76	0.24	0.06
NNSA	4	0.24	0.76	0.24	0.06
NNSA	5	0.22	0.78	0.22	0.06
NS	1	0.76	0.24	0.76	0.06
NS	2	0.80	0.20	0.80	0.06
NS	3	0.60	0.40	0.60	0.07
NS	4	0.36	0.64	0.36	0.07
NS	5	0.18	0.82	0.18	0.05

C.5.1.2 /ʌ-ɛ/ continuum

Table C.49: Analysis of deviance, L2-label categorisation task: /ʌ-ɛ/ continuum

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			749	1017.07		
Group	2	7.88	747	1009.20	0.0195	*
stim	4	295.48	743	713.71	< 2.2e-16	***
Group:stim	8	47.80	735	665.92	1.081e-07	***

Table C.50: GLM results, L2-label categorisation task: /ʌ-ɛ/ continuum. Interaction between group and stimulus as predictor.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-2.3273	0.5238	-4.44	8.88e-06	***
GroupNNSB	1.2528	0.6084	2.06	0.0395	*
GroupNS	-0.4243	0.7931	-0.53	0.5927	
stimvtef2	2.0136	0.6046	3.33	0.0009	***
stimvtef3	4.0190	0.6660	6.03	1.60e-09	***
stimvtef4	5.3953	0.8931	6.04	1.53e-09	***
stimvtef5	19.8933	589.7528	0.03	0.9731	
GroupNNSB:stimvtef2	-1.1950	0.7316	-1.63	0.1024	
GroupNS:stimvtef2	-0.7784	0.9250	-0.84	0.4000	
GroupNNSB:stimvtef3	-2.3848	0.7861	-3.03	0.0024	**
GroupNS:stimvtef3	-0.6041	0.9420	-0.64	0.5213	
GroupNNSB:stimvtef4	-3.2463	0.9946	-3.26	0.0011	**
GroupNS:stimvtef4	-0.9856	1.1406	-0.86	0.3876	
GroupNNSB:stimvtef5	-17.7443	589.7529	-0.03	0.9760	
GroupNS:stimvtef5	-13.2500	589.7539	-0.02	0.9821	

Table C.51: GLM Prediction: /ɛ/ (*bed*) responses for each /ʌ - ɛ/ continuum step. Observed proportions for each response, fitted values, and standard error.

Group	stim	cup	bed	fitted	se
NNSB	1	0.75	0.25	0.25	0.06
NNSB	2	0.56	0.44	0.44	0.07
NNSB	3	0.36	0.64	0.64	0.06
NNSB	4	0.25	0.75	0.75	0.06
NNSB	5	0.25	0.75	0.75	0.06
NNSA	1	0.91	0.09	0.09	0.04
NNSA	2	0.58	0.42	0.42	0.07
NNSA	3	0.16	0.84	0.84	0.05
NNSA	4	0.04	0.96	0.96	0.03
NNSA	5	0.00	1.00	1.00	0.00
NS	1	0.94	0.06	0.06	0.03
NS	2	0.82	0.18	0.18	0.05
NS	3	0.34	0.66	0.66	0.07
NS	4	0.16	0.84	0.84	0.05
NS	5	0.02	0.98	0.98	0.02

C.5.2 L1-label categorisation task

Table C.52: Analysis of deviance, L1-label categorisation task, /ɑ-ʌ/ continuum.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			549	411.52		
stim	4	35.26	545	376.26	4.115e-07	***
Group	1	0.65	544	375.61	0.4198	
stim:Group	4	2.58	540	373.03	0.6302	

Table C.53: GLM results, L1-label categorisation task, /ɑ-ʌ/ continuum. Stimulus as single predictor.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-2.3026	0.3317	-6.94	3.85e-12	***
stimasvt2L1	-0.9746	0.6078	-1.60	0.1088	
stimasvt3L1	-0.1153	0.4806	-0.24	0.8104	
stimasvt4L1	0.2928	0.4441	0.66	0.5097	
stimasvt5L1	1.4116	0.3925	3.60	0.0003	***

Table C.54: GLM Prediction: /o/ (*ron*) responses for each /ɑ - ʌ/ continuum step. Observed proportions for each response, fitted values, and standard error.

Group	stim	pan	ron	fitted	se
NNSB	1	0.93	0.07	0.09	0.03
NNSA	1	0.98	0.02	0.09	0.03
NNSB	2	0.89	0.11	0.04	0.02
NNSA	2	0.91	0.09	0.04	0.02
NNSB	3	0.62	0.38	0.08	0.03
NNSA	3	0.89	0.11	0.08	0.03
NNSB	4	0.96	0.04	0.12	0.03
NNSA	4	0.95	0.05	0.12	0.03
NNSB	5	0.87	0.13	0.29	0.04
NNSA	5	0.76	0.24	0.29	0.04

Table C.55: Analysis of deviance, L1-label categorisation task, /ʌ-ɛ/ continuum.

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)	
NULL			549	759.83		
stim	4	183.51	545	576.32	< 2.2e-16	***
Group	1	48.27	544	528.05	3.718e-12	***
stim:Group	4	30.46	540	497.60	3.950e-06	***

Table C.56: Logistic regression results, L1-label categorisation task, /ʌ - ɛ/ continuum. Interaction between group and stimulus as predictor.

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-1.9253	0.4046	-4.76	1.95e-06	***
GroupNNSB	0.2939	0.5446	0.54	0.5894	
stimvtef2L1	1.0343	0.5018	2.06	0.0393	
stimvtef3L1	4.2279	0.6194	6.83	8.76e-12	***
stimvtef4L1	5.2024	0.8262	6.30	3.03e-10	***
stimvtef5L1	5.9143	1.0872	5.44	5.33e-08	***
GroupNNSB:stimvtef2L1	-0.6792	0.7009	-0.97	0.3325	
GroupNNSB:stimvtef3L1	-2.7788	0.7680	-3.62	0.0003	***
GroupNNSB:stimvtef4L1	-3.1656	0.9440	-3.35	0.0008	***
GroupNNSB:stimvtef5L1	-3.7232	1.1804	-3.15	0.0016	**

Table C.57: GLM Prediction: /e/ (*red*) responses for each /ʌ - ɛ/ continuum step. Observed proportions for each response, fitted values, and standard error.

Group	stim	pan	ron	fitted	se
NNSB	1	0.89	0.11	0.16	0.05
NNSA	1	0.73	0.27	0.13	0.04
NNSB	2	0.11	0.89	0.22	0.06
NNSA	2	0.02	0.98	0.29	0.06
NNSB	3	0.00	1.00	0.45	0.07
NNSA	3	0.84	0.16	0.91	0.04
NNSB	4	0.78	0.22	0.60	0.07
NNSA	4	0.55	0.45	0.96	0.03
NNSB	5	0.40	0.60	0.64	0.06
NNSA	5	0.36	0.64	0.98	0.02