Università
Ca'Foscari
Venezia

Master's Degree Programme

in Scienze del Linguaggio

Linguaggio e Cognizione

Final Thesis

The effect of L1 variety on the perception of interdental fricatives by Italian learners of L2 English

Supervisor
Pavel Duryagin

Graduand
Alessia Vignoli
893467

Academic Year
2023/2024

Abstract

The study aimed to examine the perception of the English interdental fricatives /θ/ and /ð/ Italian learners of L2 English. The two sounds are allophones in Florentine Italian but not in Standard Italian. The primary objective of the study was to investigate whether the Florentine phonological inventory has an impact on the perception of these sounds. Subsequently, the research had two specific goals. First, to assess the extent to which perceived similarity of second language sounds predicts the accuracy of identifying non-native consonants. This was measured through a within-language perceptual similarity task. Second, to explore whether the phonological inventories of the listeners' first language played a role in influencing perception and identification accuracy. In order to gain insights into potential learning difficulties, the study employed the Speech Learning Model theory. By doing so, it aimed to further understand the underlying processes associated with acquiring the target sounds.

Two groups of native Italian listeners, differing in dialect, participated in the study: Florentine speakers and non-Tuscan speakers. In Experiment 1, the participants were asked to rate the similarity of selected pairs of English sounds. The results of this experiment provided predictions for the identification accuracy of these English sounds, which were then examined in Experiment 2. Overall, the outcomes of the study demonstrated that the phonological inventory of an individual can significantly influence their perception of the target sounds /θ/ and /ð/. It is important to note, however, that perceived similarity did not predict identification accuracy, as both cohorts exhibited a comparable distribution of errors.

Table of Contents

Introduction

From a purely linguistic perspective, each human being differs from the others in the way they use the language, but also in the way they perceive languages, and even more specifically speech sounds. The complex field that deals with the perception of speech sounds is Speech Perception. Its development started in the 1950s, and since then it has been evolving and becoming more and more specific.

Of particular interest for my paper is the concept of cross-language speech perception and its demonstration of how the native language influences the perception of speech sounds in a foreign language. This led me to ponder the applicability of this phenomenon to L1 Italian learners of English, a group that closely aligns with my own experiences. Thus, I embarked on a quest to uncover the underlying reasons behind the difficulties encountered by non-native speakers in perceiving unfamiliar sounds in a second language. I adopted the perspective that the challenges in L2 sound perception and production extend beyond merely L1, but delve deeper into the realm of the psycholinguistics.

Given my fascination with the topic of cross-language similarity, I decided to focus my research on the perception issues faced by Italian native speakers when it comes to English as a second language (L2). Specifically, my attention centred on two English sounds: /θ/ and /ð/. I observed that these sounds are frequently mispronounced by ordinary Italian speakers, prompting me to consider them as compelling subjects for investigation.

Furthermore, the relationship between these two English sounds and the Italian language intrigued me. While these sounds are not part of Standard Italian phonetic inventory, they are indeed present in Florentine, a dialect spoken in Florence. Florentine speakers produce these sounds in specific positions within words as allophones, which are in complementary distribution with the sounds /t/ and /d/ that are included in the Standard Italian inventory. This means that certain phonemes in Florentine are expressed as different variant forms (i.e., allophones), and it is possible that not all speakers of a particular L1 use these allophones. This phenomenon -later referred to as *gorgia toscana*- is evident among Florentine speakers compared to Standard Italian speakers (and also non-Tuscan speakers in general), as the former possess and employ the allophones [θ] and [ð] that are not present in Standard Italian. It is worth noting that /θ/ and /ð/ are conversely phonetically meaningful in English and can alter the meaning of words.

Consequently, in my current study, I aim to investigate how Florentine learners of English perceive the English sounds /θ/ and /ð/. For these learners, these sounds function as allophones of the Standard Italian sounds /t/ and /d/, which is not the case for non-Tuscan speakers. To ensure accuracy and avoid any confounding factors, I have divided the participant groups into Florentine and non-Tuscan categories. By doing so, I hope to account for any potential overlaps where Tuscan speakers (who are not Florentine) may exhibit Florentine accent features to a certain extent.

The primary objective of this study is to investigate the extent to which Italian learners of English perceive the sounds /θ/ and /ð/ differently from /t/ and /d/, that are allophones in Florentine Italian. Starting from there, I want to check if their perception differs from non-Tuscan learners of English who do not have /θ/ and /ð/ as allophones in their language. Specifically, I am interested in Italian speakers who also speak Florentine and whether they encounter difficulties in perceiving these sounds, which could potentially impact their ability to learn them. Furthermore, I aim to compare their perception to that of non-Tuscan speakers who do not produce these sounds when speaking Italian.

To make predictions about these potential differences, I rely on the Speech Learning Model (SLM), a second language learning model first proposed by Flege (1995) and later revised by Flege and Bohn (2021). According to SLM, L1 and L2 categories share the same perceptual space, but various methods exist for assessing how L2 sounds are perceived by L1 listeners. These methods examine the perceptual relationship between native and non-native sounds in the mind of L1 listeners, who are learning an L2, and whether any issues may arise from this relationship. Some scholars suggest that direct measures are the most reliable methods for assessing the L1-L2 relationship and its consequences on the perception of the sounds (Bohn, 2002; Strange, 2007). Following this recommendation, my research employs two approaches aligned with Speech Perception studies.

Firstly, in Experiment 1, listeners are presented with two auditory stimuli and asked to provide graded ratings of the perceptual similarity between the two non-native speech sounds. This task is known as a perceptual similarity or within-language graded discrimination task (Bohn, 2002; Bohn & Ellegaard, 2019). Secondly, in Experiment 2, listeners are required to provide a discrete orthographic label for a non-native stimulus from a closed set of forced-choice alternatives. This task is referred to as an identification task (Bohn & Flege, 1992; Bohn &Steinlen, 2003).The experiments serve as the fundamental component of my research and are outlined in Chapter 3. The paper, however, has been organised in a manner that first establishes a theoretical foundation before delving into the empirical aspects. Its structure can be presented as follows.

Chapter 1 provides an in-depth exploration of the Florentine dialect from a phonological standpoint, specifically examining its contrasts with Standard Italian and the English language, as well as distinguishing features from other Tuscan dialects. The chapter first offers a historical account of the dialect's development and its geographical distribution, having strong influence on the development of Standard Italian language. Subsequently, an analysis of the dialect's phonological inventory is presented, followed by a comparative examination with the phonological inventory of English, with particular attention given to the target sounds /θ/ and /ð/ and to the so-called “gorgia toscana”. *Gorgia* is an ongoing process in Tuscan dialects, and it originated in Florence and spread to surrounding areas. It is responsible for the pronunciation of specific phonemes, namely /p/, /t/, and /k/, as [ɸ], [θ], and [h/x] respectively. This redefinition introduces a process known as *spirantisation*, wherein the occurrence of spirantised forms alternates with complete occlusion attested after a pause or after a consonant.

In Chapter 2, the research delves into the broader field of Speech Perception. The historical background is initially provided, starting from the 1950s up until 1995, when the Speech Learning Model (SLM) was theorised by Flege. The SLM is based on the perception of phonetic differences between L1 and L2 sounds, and categorises L2 sounds based on their similarity to corresponding native language sounds, using three categories: "new," "similar," and "identical." It suggests that perception difficulties experienced by L1 listeners when exposed to nonnative sounds are influenced by the classification of the L2 sound into those categories. It is therefore explained why this theoretical framework serves as the foundation for the present study and stands as a reference point for the final analysis. SLM also introduces the relationship between L1 and L2 phonological systems, known as "similarity." The chapter elucidates the methodologies employed in Speech Perception studies that are linked to the SLM. Finally, the chapter anticipates the two types of experiments that form the core of my investigation, namely within-language graded discrimination task and identification task.

Chapter 3 delves into the exact research methodology, presenting the two experiments and reporting their respective data. The study predicts that native Florentine listeners will perceive these pairs as very similar or identical, compared to non-Tuscan listeners who are expected to perceive a higher level of dissimilarity. The impact of dialect on contrasts may vary, with Florentine listeners perceiving /ð/ and /d/ as less similar compared to voiceless counterparts. The dissimilarity ratings provide insights into the predicted accuracy of sound identification. Experiment 1 is a within language perceptual similarity test, whose results are used to predict identification accuracy, which in turn is tested in Experiment 2. Two separate sections are dedicated to each experiment, but they have the same structure. Moreover, they share the groups of participants. Data analysis is both descriptive and statistical. Statistical analysis is based on generalised linear mixed-effects regression models (GLMMs) with the aid of R package *lme4* (Bates et al., 2015), and whose results are commented on with the integration of the Language Profile Questionnaire responses (LPQ).

Chapter 4 summarises the results, comprehensive of both experiments. It is mainly a discussion of the results, where I present the results, compare them to the research questions, speculate about possible explanations, and add interpretations. The results showed that Florentines generally had higher similarity ratings for English initial interdental fricative consonants, but the identification patterns of target sounds were nearly identical across both groups. Four hypotheses were tested to explore perceived similarity ratings and potential between-group differences, as well as their impact on L2 identification accuracy. The results only partially supported H1, suggesting that the influence of dialect would vary depending on voicing. The study did not support H4, which pertains to the identification of target sounds.

Chapter 5 concludes with the main points and goals achieved. It underlines that the study's main focus is on the perception of English sounds by speakers from Florence, specifically on the influence that a different phonological inventory -as is Florentine, which uses [θ] and [ð] as allophones of /t/ and /d/- can have on identification patterns. The primary result is that Florentine speakers perceived target sounds more similar than non-Tuscans, particularly for the voiced interdental fricative. Nonetheless, the perceived similarity test predicted identification accuracy to a limited extent. This leads to the conclusion that the phonological inventory did not affect correct identification, and this did not depend on dialectal properties. In conclusion I also acknowledge that the identification of certain sounds might be more dependent on the L1 inventory and exposure to English, rather than the variety of phonological inventories, and acoustics properties.

The final chapter concludes with a set of limitations concerning the study. Among these, the nature and the structure of the stimuli, the number of data collected, connected to the number of participants reached and the selection criteria used. Further investigation is recommended to validate the interpretation. In spite of the drawbacks, it is possible to acknowledge that this was the first attempt to focus on a specific dialect using an underexplored method as is the similarity task, and some concrete conclusions have been reached.

1. Phonological inventories

1.1 The Florentine dialect

The Tuscan dialect of Florence, commonly referred to as Florentine, held official status and exerted a significant influence on the development of Standard Italian language (Bertinetto&Loporcaro, 2005). It achieved literary recognition in Italy during the period from the 14th to the 16th century[1]. The expansion of Florentine beyond Tuscany can be attributed to a group of educated individuals who embraced it for academic and administrative purposes. Subsequently, with the establishment of the school system in the post-unification era, it further proliferated (Calamai, 2017). According to De Mauro (1972), Florentine officially began to be used for everyday communication all over the Country in the 20th century. From a linguistic standpoint, Standard Italian gradually absorbed the phonetic characteristics of Florentine. However, scholars note that this is not the case for some allophonic processes (Bertinetto&Loporcaro, 2005; Calamai, 2017). This distinction is of utmost importance for the purpose of this paper, as elaborated below.

1.1.1 Florentine vs. Standard Italian

Although Florentine serves as the foundation for the origin of Standard Italian, not all of its features can be attributed to Florentine alone. Bertinetto and Loporcaro (2005) conducted an experiment to examine the pronunciation of Italian among native speakers of different dialects, specifically from the regions of Florence (FI), Rome (RI), and Milan (MI). They compared these dialects to the formal transcription of Standard Italian (SI) in order to identify any discrepancies. The results revealed that FI and RI closely resemble SI (which, as

[1] "The variety of Italian spoken in Florence is the result of the uninterrupted spontaneous evolution of the language [...] and this indirectly extends to Rome, since Roman speech was heavily tuscanised in the late Middle Ages. [...] The original Italo-Romance vernacular, which shared many isoglosses with southern Italian dialects, was ousted by a tuscanised one" (Bertinetto&Loporcaro, 2005, pp. 144-145).

mentioned earlier, is deemed to be Florentine-based) only in formal contexts, when individuals are aware of being recorded. In informal situations, however, FI may exhibit variations compared to SI. Calamai (2017) further explains that the phonetics and phonology of casual speech differ between Standard Italian and Florentine. Distinctive features that set Florentine apart from Standard Italian include the spirantization of intervocalic voiceless stops, consonantal weakening (known as "Gorgia Toscana"), s-affrication, stressed vowel systems, and *raddoppiamentosintattico*[2] (Calamai, 2017).

1.1.2 Florentine vs. Tuscan dialects

From a historical standpoint, Florentine can be considered a dialect (Giacomelli, 1975, as cited by Calamai, 2017), albeit one among the numerous Tuscan dialect variants in use today. The Tuscan dialects are geographically demarcated by the "La-Spezia Rimini line" isoglosses, making it relatively straightforward to distinguish Tuscan dialects from non-Tuscan dialects, starting from the northern region of Tuscany. However, identifying Tuscan features becomes more challenging along the southern and eastern borders due to the contact between Tuscany, Umbria, and Latium (Calamai, 2017, p. 216).

When considering the various language variations in Tuscany, it is common to divide the region into six dialects: Florentine, Sienese, western Tuscan (further divided into Pisan, Leghornese, Elbano), Aretino, Grossetano-Amiatino, and Apuano (Calamai, 2017, p. 216). However, an exception arises with the work of Giannelli (2000), which identifies ten different varieties. The reason for this difference is that Giannelli's classification is based on more detailed analysis of morphological and syntactical features. Despite these variations, all descriptions agree that the Florentine dialect stands on its own and has had a significant

[2] A phonological process whereby word-initial consonants undergo sandhi gemination after final stressed vowels when no pause intervenes. Bertinetto and Loporcaro (2005, p. 135) provide the following examples: “tre case [tre 'k:a:se] 'three houses', parlòlatino [parlo l:a'ti:no] 's/he spoke Latin'”. It applies further to a small set of weak monosyllables and paroxytones such in the case of “a te [a 'tie] 'to you', come voi ['kołmeVioi] 'like you.pl'” (Bertinetto&Loporcaro, 2005, p.135).

influence on other dialects, particularly in terms of the spirantization of stops, also known as the "Gorgia" (Calamai, 2017, p. 219), which has become the most distinctive feature of Tuscan Italian.

1.2 Gorgia toscana

While the phoneme inventory of Standard Italian closely resembles that of the Florentine dialect, there are certain allophonic phenomena absent in Standard Italian. One prominent feature, and the main focus of this paper, is the *gorgia toscana*or *Tuscan gorgia*. Historically, it is believed that the *gorgia* existed during Dante's time (late 13th/early 14th century), and Florence is considered as its epicentre. It is documented that the *gorgia* quickly spread to the surrounding areas and became a characteristic feature of Tuscan pronunciation. However, the earliest evidence of the *gorgia* dates back to the mid-16th century with Claudio Tolomei, a Siena-born grammarian, who published "Polito"[3], demonstrating the aspiration of /k/ in the *gorgia toscana*.

Scholars have used various terms to describe this process in literature. Bertinetto and Loporcaro (2005) refer to it as "spirantization", Calamai (2017) as "consonantal weakening", and Marotta (2008) as "lenition process"[4] or the literal translation "Tuscan throat". Marotta (2008) explains that "lenition" generally describes a process of weakening and that it can refer to different processes depending on two different Italian traditions. One tradition involves stops transforming into fricatives, affricates, or approximants (Lass, 1984, p. 178), while the other tradition involves a partial voicing of stops, commonly found in central and southern dialects. Marotta employs the former interpretation when discussing the *gorgia toscana*.

[3]A book published in 1525 by Tolomei, by which he reformed Italian orthography.
[4]Marotta (2008) uses this term to indicate the sound that is meant to be produced. In the case of *gorgia* this would be /t/ and /d/. It is used as opposed to "output" sounds, which would be the actual realisations of those sounds, the allophones [θ] and [ð].

1.2.1 Triggering and blocking contexts

The application of *gorgia* in Tuscan dialect targets both voiced and voiceless stops, as well as fricatives and sonorants to a lesser extent. Calamai (2017, p. 219) provides a more detailed description of this phenomenon, as follows:

> *[Gorgia Toscana is] Responsible for the pronunciation of /p/ /t/ /k/ as [ɸ] [ɵ] [h/x], with the plosive preceding the sonorants [r], [l], or the approximants [j], [w] and only following vowels or glides, both internal to and across words: la [h]asa la casa 'the house', la [h]rema la crema 'the cream', la [h]uercia la quercia 'the oak'. The spirantized forms alternate with full occlusion in the case of post-pause – [k]asa, [k]rema, [k]uercia – or following a consonant – in [k]asa 'in the home', con [k]rema 'with cream', con [k]uercia 'with oak' – or in case of raddoppiamentosintattico: tre [kk]ase 'three houses', tre [kk]reme 'three creams', tre [kk]uerce 'three oaks'. To a lesser extent, the corresponding voiced plosives /b/ /d/ /g/ are also affected by spirantization and variable realized as [β] [δ] [ɣ] in intervocalic and postvocalic position. [...] Voiceless phonemes appear to be more sensitive to the weakening process than voiced ones and, among the voiceless class, velar segments are more affected than dentals and labials.*

Bertinetto and Loporcaro (2005) concur with Calamai's view that *gorgia toscana* is influenced by the place of articulation: velar stops are more susceptible to lenition than dental and labial stops. It has been observed that *gorgia* can occur under specific conditions determined by syllabic rules. Specifically, Tuscan lenition occurs only after a vowel or semivowel, and for input stops, lenition occurs when they precede a vowel, glide, or liquid. Furthermore, a morphosyntactic analysis reveals that *gorgia toscana* can be applied across phrase and word boundaries. Conversely, lenition is said to be blocked when the stop is preceded by a "heterosyllabic consonant or a pause" (Marotta, 2008, p. 235).

It is important to note that there is no contrast between stops and their spirantised counterparts, namely [θ]-[t] and [ð]-[d], which are considered allophones in the literature. In other words, both word-internally and word-initially, these two classes of segments function as allophones and exhibit complementary distribution.

To illustrate this, Table 1 below compares the consonant system of Standard Italian to that of Florentine, focusing on the sounds relevant to the study that are associated with *gorgia*. The sounds enclosed in square brackets [] represent the allophones in the Florentine inventory. Therefore, all the sounds listed in the table are part of the Florentine inventory, while only the phonemes enclosed in slashes // belong to Standard Italian.

	Bilabial	*Dental*	*Alveolar*	*Velar*	*Glottal*
Plosive	/p/ /b/		/t/ /d/	/k/ /g/	
Fricative	[ɸ] [β]	[θ] [ð]		[x] [ɣ]	[h][5]

Table 1

The Florentine phonological inventory restricted to plosives and their fricative allophones

1.2.2 Geographical distribution of gorgia

Research conducted by Giannelli and Savoia (1978, 1979-80) suggests that *gorgia* originated in Florence and then spread to the surrounding areas. However, it is important to note that *gorgia* is an ongoing process that continues to evolve, and not all Tuscan varieties are equally affected by it (Marotta, 2008, p. 240). Some varieties, such as those in Massa, Carrara, and Northern Lucca, do not exhibit *gorgia toscana.* Additionally, in the Eastern area of Tuscany, which includes Arezzo, *gorgia* is not consistently present. It occurs only with the /k/ sound and conflicts with another form of consonantal weakening found in other varieties of Central

[5] Note that there is a difference in distribution as for the sounds [x] and [h]. They are both allophones of the Standard Italian phoneme /k/, influenced by the same context as the one described for /t/ and /d/. However, the different use of [x] and [h] has geographical reasons. Specifically, some studies have revealed that Pisan Italian always shows [x] as allophone for /k/, whereas in Florentine Italian [h] is the most frequent, while [x] occurs only in back vowel contexts (Calamai, 2017, p. 220).

and Southern Italy. On the contrary, *gorgia* is active in the areas surrounding Florence, Siena, Prato, and Pistoia, with Florence considered the epicentre of *gorgia toscana*. Therefore, *gorgia toscana* is not uniformly distributed across Tuscany, contrary to what the name might suggest. It originated in Florence and spread throughout Tuscany, but not to the same extent in all areas. It can only be attributed to the specific areas mentioned above.

Variability is also observed when considering the quality of the lenited stop. Casentino and southern Tuscany, which occupy peripheral positions in the region, serve as examples. In these areas, intervocalic /p/ /t/ /k/ are produced as lenited[6] [p̥ t̥ k̥] or as fully voiced variants [b d g], following the same distribution patterns in the speech as *gorgia toscana* (or *Tuscan gorgia*) typical of Florence (Calamai, 2017, p. 221).

Calamai (2017) highlights that Tuscan speakers are generally aware of their dialectal features, with /k/ lenition ("aspirazionedella c") being the most commonly recognized. In fact, Tuscan people (and, generally, Italian people) refer to *gorgia* as "aspirazionedella c" (referring to the orthographic representation of /k/), which might lead to the observation that they are only aware of the process for the velar point of articulation[7] (although some participants in my study acknowledged awareness of "aspirazione" of /t/ and /d/ (see Appendix A). It is not just Tuscan people who are aware of this phenomenon. On the contrary, the spirantisation of stops, particularly /k/, is associated with being Tuscan. In this sense, *gorgia* holds social significance as it serves as a marker of local identity.

1.2.3 Acoustics of gorgia

Despite extensive literature on the acoustic analysis of *gorgia*, the discussion is still ongoing (among others, Giannelli & Savoia, 1978; Marotta, 2008). Marotta (2008, p. 239) describes

---

[6] A more sonorous plosive with weakened articulation, but not a fricative in the proper meaning.
[7] However, some participants in my study acknowledged awareness of "aspirazione" of /t/ and /d/ when speaking or talking to other people, according to the questions of the LPQ (see Appendix A).

the spectrograms of lenited consonants with the following acoustic features: a slight voicing at low frequencies, lower amplitude compared to voiced stops, the presence of friction noise, and a lack of burst or very low energy level. Additionally, lenited stops sound and appear longer than their voiced counterparts but shorter than voiceless stops.

Several recent studies have highlighted the systematic differences in acoustic realisation. This is particularly evident in the case of Florentine, Pisan, and Sienese dialects, as stated by Calamai (2017, p. 220): "For instance, in Pisan Italian, /k/ is most frequently realised as [x], regardless of the surrounding vowel context, whereas in Florentine Italian [x] was found to occur only in back vowel contexts and [ɦ] occurred most frequently elsewhere (Sorianello 2001, 2003). From this respect, Sienese Italian seems to be an intermediate between the two, having a relatively high frequency of realisation of [x] (Stevens 2012 contra Giannelli 2000)".

It is important to note that *gorgia* and its realisation are not uniform. Due to the variability observed in Tuscan dialects, especially in relation to *gorgia*, my analysis focuses solely on the perception of Florentine speakers from the province of Florence. Additionally, I specifically recruited non-Tuscan individuals for the comparison group. This approach aims to minimise potential confounding factors and to avoid the need to address differences among speakers from various regions in Tuscany, who may only share certain linguistic features, as discussed above. A more detailed discussion of this matter will be provided in Chapter 3 when presenting the study.

1.3 Florentine vs. English

According to Wheelock (2016), the complete phonemic inventories of Standard Italian and English are similar. However, for the purpose of this paper, the focus is on fricatives (and possibly the plosives /t, d/). Consequently, it can be observed that Italian lacks /θ, ð, h/ in comparison to English fricatives (Wheelock, 2016). Florentine, on the other hand, displays different behaviour in this regard. When considering the phonological inventory, including allophones, Florentine exhibits more similarities to the English inventory. This can be seen in the table below (Table 2), which illustrates the place of articulation for plosives and fricatives in both English and Florentine. Common sounds are denoted in black, Florentine-specific sounds in violet, and English-specific sounds in blue.

	Bilabial	*Labiodental*	*Dental*	*Alveolar*	*Post-alveolar*	*Velar*	*Glottal*
Plosive	/p/ /b/		/t/ /d/	/t/ /d/		/k/ /g/	
Fricative	[ɸ] [β]	/f/ /v/	/θ/ /ð/ [θ] [ð]	/s/ /z/	/ʃ/ /ʒ/ [ʒ]	[x] [ɣ]	/h/ [h]

Table 2

Florentine and English phonological inventories as for plosives and fricatives. Note: violet is used for the Florentine-specific sounds; blue is used for the English-specific sounds; black indicates the shared (and ideally identical) sounds.

It is evident that the English and Florentine languages share certain plosive sounds. However, there is a distinction in the place of articulation for the /t/ and /d/ sounds. In English, these sounds are realised as alveolar sounds, while in Florentine (similar to Standard Italian), they are pronounced closer to dental sounds. Voice Onset Time (VOT) is another important difference between the two languages. Some scholars suggest that these differences together may lead to mispronunciation of English sounds (Wheelock, 2016). This aspect could be relevant for discussion in Chapter 4.

Regarding fricatives, both Florentine and English have the same voiced and voiceless labiodental sounds /f, v/, alveolar sounds /s, z/, and postalveolar sounds /ʃ, ʒ/. They also share the dental sounds /θ, ð/, but with a significant difference. In Florentine, these sounds serve as allophones of /t, d/ and do not change the meaning. In contrast, they are phonemes in English. The same distinction applies to the sound /h/[8]. It is important to note that the allophones in Florentine result from the lenition of plosive sounds, known as *gorgia toscana*, as discussed in 1.1.3 above. This study aims to specifically examine the perception of /θ/ and /ð/ by Florentine speakers, who use them as allophones, in comparison to Standard Italian speakers who do not have /θ/ and /ð/ in their phonetic repertoire and do not pronounce them in their native language.

1.3.1 A note on /θ/ and /ð/

Given that my research question focuses on /θ/ and /ð/, it is necessary to emphasise some important points regarding these sounds in the context of Italian native speakers acquiring English as a second language.

First and foremost, these sounds are not part of the standard Italian phonetic inventory (Berti 2015) and are often substituted with /t/ and /d/ (Avery & Ehrlich, 1992, p. 132). It is crucial to note that they are indeed present in the phonological repertoire of Florentine. As previously mentioned, they serve as allophones of /t/ and /d/ in the Florentine dialect. According to the literature, /t/ and /d/ undergo a lenition process in specific phonological contexts, resulting in the production of /t/ and /d/ as their fricative counterparts, namely [θ] and [ð] (also known as *gorgia toscana*).

[8] This is indicated by means of different notation in the table. According to IPA rules, // denotes phonemes, [] are used in case of allophones.

2.Second Language Speech Perception

The general perception of language is believed to be based on the interpretation of speech sounds that are associated with the phonemic categories of our first language (L1), which we acquire from birth. This viewpoint is supported by scholars such as Best (1995) and Flege (1995) in relation to L1. Consequently, individuals from different language backgrounds may perceive the same stimulus differently, depending on the development of sound categorization that occurs during language acquisition (specifically in relation to attending to specific acoustic aspects of speech) and is specific to their language (Schmidt, 2007). This concept has also been applied to studies on the perception of second language (L2) or target language (TL), further supporting the assertion that one's L1 can influence the perception and subsequent acquisition of L2 (Flege, 1995; Schmidt, 2007), as will be discussed later in this chapter.

According to contemporary researchers, the study of L2 acquisition emerged as a recognised scientific field in the mid-20th century. Through scientific methodologies, robust models of L2 acquisition have been developed (Thomas, 2013). Most scholars today acknowledge that a learner's L1 plays a role in L2 acquisition, although the nature of this role is a subject of debate and a distinguishing factor among different models. The subsequent section discusses the development of these models, focusing on the most frequently cited ones, in order to establish the foundation of the core theory that my study is based on, namely the Speech Learning Model (SLM) by Flege (referred to as SLM).

2.1 Brief overview of L2 Speech Perception theories prior to SLM

2.1.1 From 1950s to 1980s

In the 1950s, when research on L2 acquisition began to develop, the initial theories aimed to explain the mechanisms of acquiring L2 sounds by comparing the phonological systems of two languages. One of the earliest approaches was the Contrastive Analysis Hypothesis (CAH) proposed by Lado in 1957 (Munro & Bohn, 2007). The CAH states that learners tend to assume that the characteristics of their first language (L1) also apply to the L2, leading to the significant influence of L1 on the acquisition of L2. Consequently, it was predicted that learning L2 phonemes without a corresponding counterpart in the L1 would be challenging, while learning L2 phonemes with counterparts would be relatively easier. The CAH focused on articulation or production difficulties rather than considering the possibility that differences in perception could affect the accuracy of target sounds. Interestingly, the CAH overlooked the significant phonetic differences that might exist between seemingly "similar" sounds in two languages, an aspect that modern theories, presented in sections 2.1.2 and 2.2, emphasise.

In 1966, a novel perspective on the topic emerged. Specifically, it was emphasised the need for a comprehensive analysis of speech sounds to gain insights into the challenges faced by second language (L2) learners. This approach, known as Error Analysis, suggested that phenomena such as "interference" and "transfer" from the learners' first language (L1) played a role in L2 acquisition[9]. However, both Error Analysis and Contrastive Analysis Hypothesis (CAH) were deemed insufficient as they could only account for a limited range of difficulties.

[9] “In second-language acquisition, the tendency to transfer the phonology, syntax, and semantics of the native language into the learning of the second language. Negative transfer (or interference) occurs when differences between the two languages’ structures lead to systematic errors in the learning of the second language or to fossilisation” (taken from APA Dictionary of Psychology, accessed December, 15th 2023).

A case in point exemplifying the integration of Contrastive Analysis and Error Analysis is the study conducted by Wheelock in 2016. Wheelock compared the consonant and vowel systems of Italian and English and employed both Contrastive Analysis and Error Analysis methodologies. It was observed that speakers of Standard Italian (SI), with a smaller set of vowel sounds compared to North American English (NAE) (7 vs. 12), faced challenges in acquiring NAE sounds. The same trend was noticed for consonants. Based solely on a comparison of phoneme inventories[10] Wheelock predicted that Italian speakers learning English as an L2 would encounter difficulties with specific sounds, such as the English dental fricatives /θ/ and /ð/, which are absent in Italian. Furthermore, Italian learners were generally observed to struggle with tongue placement and the appropriate aspiration[11].

Over time, this approach has been further developed, significantly influencing the field of speech perception research and reframing core theoretical questions that have prevailed since the 1950s (Pisoni, 2018). I will now proceed to present these influential advancements.

2.1.2 From the 1980s to 1995

The prevailing notion, established in the 1950s, that a learner's native language (L1) has some influence on the acquisition of a second language (L2) was supported by research conducted in the 1980s. However, this viewpoint was examined from a new perspective.

[10]Even though the methodology used by Wheelock (2016) is different from mine, and it is also based on basic assumptions, some of her considerations ought to be taken into account.

[11]As commented by Wheelock (2016, 52), "they tend to keep the tongue inside and press it against the teeth, forming the alveolar stop [t] or placing the teeth on the lip, forming the labial-dental fricative [f] (for /θ/) [...]. Similarly, they also tend to form the alveolar stop [d] (for /ð/)". The explanation given for the misformation of the dental fricative /ð/ as the alveolar stop [d] is that the former is not part of the Italian inventory, and the Italian [d] is the closest phoneme to it. The same reason is provided for the English voiceless dental fricative /θ/, mostly substituted with the alveolar stop /t/.

Thomas (2013) references White's (1985) and other studies carried out in the 1980s, proposing a different explanation for the impact of L1 on L2. It suggests that the influence of L1 on L2 does not stem from L1 sounds, words, or sentences serving as models for their L2 counterparts (as previously believed). Rather, it posits that the learners' understanding of the language underlying the units and their structured distribution in L1 is the source of influence (or lack thereof) on L2. This prompted further research to consider the linguistic context, such as allophones, and to analyse L2 speech learning through a different lens: perception (Schmidt, 2007). This led to the emergence of studies in speech perception, specifically in relation to L2 (or cross-language study).

L2 perception study is a subset of the broader field of L2 Speech Perception, which encompasses phonetics, phonology, cognitive psychology, L2 acquisition, and applied linguistics. Its primary concern is how L2 speakers process L2 speech sounds, with the ultimate goal of explaining the mechanisms that contribute to the development of L2 speech (Chang, 2019). Numerous approaches can be employed to investigate this phenomenon, but the phonological and phonetic aspects are more commonly explored[12].

According to Schmidt (2007, p. 186), L2 Speech Perception can be defined as the analysis of the perceptual processing of acoustic or gestural information by a listener who already possesses a specific system of sounds, known as a phonological inventory. This means that when an L1 listener perceives an L2 sound, they do so through the lens of their own L1 phonology rather than accurately perceiving the L2 sound, at least initially. The study of L2 Speech Perception aims to predict and explain the perception process of non-native sounds by examining their similarity and dissimilarity. This research follows in line with that objective.

[12]However, they are bound to individual variables, such as age, exposure, use of L2, and also motivation or aptitude (Chang, 2016).

Best and Flege are prominent figures in the field of L2 Speech Perception studies. They have introduced two influential theories, namely, Best's Perceptual Assimilation Model (PAM) and Flege's Speech Learning Model (SLM) (Bohn & Munro, 2007). Their primary focus is to explain the challenges faced by L1 listeners when perceiving L2 sounds. They specifically explored the perception of L2 phonetic distinctions by adult speakers whose L1 lacks such distinctions. Additionally, they investigated the impact of language experience on this perception (Bohn & Munro, 2007).

Best's Perceptual Assimilation Model (PAM) (Best, 1995) outlines a set of assimilation patterns that L2 learners rely on when perceiving a sound. Consequently, the perception and discrimination of specific L2 sounds are influenced by these assimilation patterns. On the other hand, Flege's Speech Learning Model (SLM) (Flege, 1995a) suggests that the perceived acoustic similarity[13] of sounds affects the acquisition of L2 sounds (Schmidt, 2007, p. 186).

This paper is built upon the foundation of Speech Perception studies and relies on Flege's SLM for data analysis and subsequent discussion. Therefore, Flege's SLM will be presented and taken into consideration for the analysis of the data in the next chapter and throughout the conclusion, in order to organise and interpret the empirical data from the experiments. It should be noted that the study of L2 Speech Perception is continuously expanding and undergoing constant changes.

[13]Similarity is a key concept in my research, given that Experiment 1 wants to measure perceived similarity and the answers to the research questions are influenced by the results of such similarity. Anyways, it is an under-explored method, as I will point out later.

2.2. The Speech Learning Model (SLM)

Flege's Speech Learning Model (SLM) was initially proposed in 1995 and subsequently revised (Flege 1995, Flege & Bohn 2021). At the core of this model lies the postulate that "the mechanisms and processes used in learning the L1 sound system remain intact over the life span" (Flege, 1995, p. 239). However, the SLM also aims to elucidate the mechanisms that can either facilitate or hinder the successful acquisition of second language (L2) sounds.

Flege begins by asserting that learners tend to categorise L2 sounds based on their similarity to corresponding sounds in their native language, a phenomenon termed "equivalence classification." Three categories are employed: "new," "similar," and "identical." According to this framework, all sounds in the L2 (or target language, TL) can be classified by the listener's perception in relation to their native language (Chang, 2019)[14].

Additionally, Flege presents a series of postulates and hypotheses pertaining to the acquisition of L2 sounds, which are extensively discussed in his various works, as in Flege (1995). For the purpose of this section, which primarily aims to introduce the SLM, hypotheses 2 and 3 are particularly relevant. They are stated as follows:

> ***H2:*** *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.*
>
> ***H3:*** *The greater the perceived phonetic dissimilarity between an L2 sound and the closest L1 sound, the more likely is that phonetic differences between the sounds will be discerned.*

(Flege, 1995, p. 239)

[14] Note that this is considered an ordinary approach to the perception of sounds by humans for any language, also their L1. What is crucial for L2 is that an inappropriate equivalence classification can lead to issues in perception and/or production of the L2 (Chang, 2019)

As already pointed out, the theory of the SLM can be visualised as the three categories along a continuum, where L2 sound classification spans from “new” to “ similar” to “identical” sounds (or vice versa). Figure 1 shows the continuum and suggests the relationship of the L2 to the L1 sound, where “new” means “less like L1” and, conversely, “identical” is “more like L1”.

Figure 1

Continuum of similarity of L2 sounds to L1 sounds (Chang, 2019, p. 434)

Flege (1995) establishes a correlation between phonological proximity, or equivalence classification, and the learnability of L2 sounds in an academic context. According to Flege's theory, the hierarchy of difficulty for learning L2 sounds is as follows: identical < new < similar. This concept is also illustrated in Figure 2, where discrepancies between L1 and L2 influences the accuracy of L2 production. It is observed that "new" and "identical" sound productions tend to be unaccented, whereas "similar" sounds can be accented[15] and therefore more challenging to acquire.

[15] “Unaccented” is the term used to indicate a native-like production of the L2 sounds, in contrast to “accented” that conveys the idea of a pronunciation that is clearly influenced by the L1.

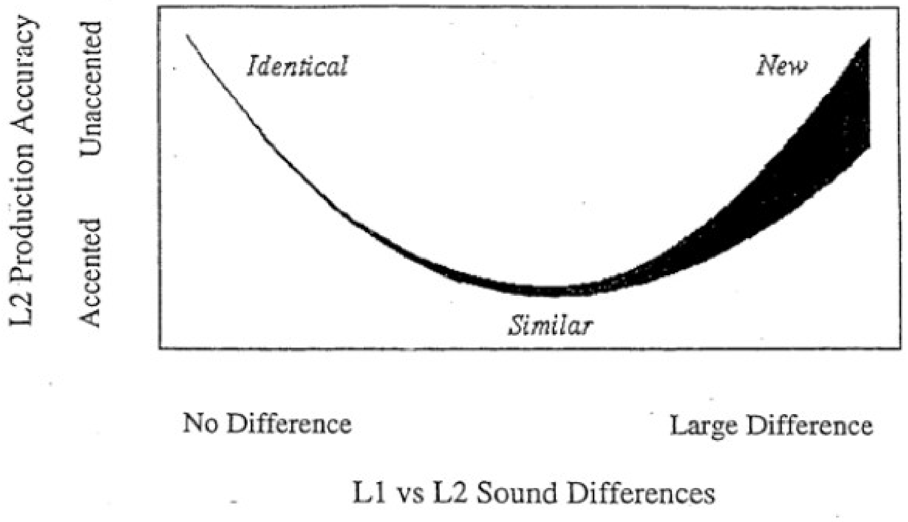

Figure 2

The continuum that represents the L1-L2 relationship in the production of L2 sounds by L2 learners. Taken from seminar's slide by Bohn, O.-S. (2023).

In summary, Flege's theory (1995) suggests that if an L2 sound is perceived as "identical," it is the easiest to learn since it shares the same characteristics as the learner's L1. "New" sounds, although still noticeably different from any L1 counterparts, present a greater level of difficulty. This is because learners tend to create a new category for these sounds, requiring them to learn additional aspects specific to the "new" L2 sound. On the other hand, if an L2 sound is perceived as "similar" and perceptually equivalent to an L1 sound, the learner fails to establish a new category for that sound. This is because "similar" sounds occupy an intermediate space of cross-linguistic similarity, as depicted in both Figure 1 and Figure 2, leading to inappropriate L1 influence on L2 sound perception and acquisition (Chang, 2019, p. 433). In conclusion, the higher the perceived cross-linguistic similarity between L1 and L2 sounds, the poorer the perception of non-native sounds. Conversely, a greater perceived phonetic dissimilarity between L1 and L2 sounds increases the likelihood of establishing new L2 categories.

Despite the challenges faced by L2 learners, Flege (1995) argues that learners have the perceptual ability to acquire non-native sounds regardless of age, throughout their lifespan. Various studies have provided evidence for the plasticity of L2 learners' perceptual systems by demonstrating that auditory training can enhance both perception and production of L2 sounds. However, this does not apply universally to all L2 sounds, as explained earlier, since it depends on the degree of cross-language phonetic similarity (Bohn & Flege, 2021). SLM has been utilised in numerous studies with diverse methodologies and objectives, but not all of them could rely on the SLM equally.

2.2.1 Studies based on SLM

The original objectives of the Speech Learning Model (SLM) were initially focused on predicting the eventual accuracy of identification and production of individual L2 sounds. However, its scope has been broadened to encompass the investigation of hypotheses related to various aspects of L2 speech perception and acquisition. For instance, SLM has been employed to test hypotheses in studies pertaining to L2 identification and subsequent discrimination (Guion et al., 2000; Bohn & Best, 2012), assimilation and discrimination (Bohn et al., 2011), and the influence of perceptual assimilation and graded discrimination on identification accuracy (Bohn & Ellegaard, 2019).

Guion et al. (2000) conducted two experiments with native Japanese speakers, focusing on the acquisition of English /r/ and /l/ by Japanese learners, as these sounds are particularly challenging to discriminate for L1 Japanese speakers. In the first experiment, participants were required to identify English and Japanese consonants based on a Japanese category and rate them according to their goodness-of-fit. These ratings were then used to predict the discrimination of L2 English sounds in a subsequent task, which involved native English speakers and three groups of Japanese speakers. The results demonstrated that the perceived phonetic distance of second language consonants did indeed predict the discrimination of second language sounds. Thus, Guion et al. (2000) confirmed the hypothesis of the Speech Learning Model, providing evidence that L2 learners are more likely to establish a phonetic category for sounds that are perceptually distant from their closest native category.

The study conducted by Bohn et al. (2011) focused on L1 Italian and L1 Danish speakers. It aimed to examine how the specific phonetic details of the phonological systems in the participants' native languages influenced their perception of three non-native English consonant contrasts: /b/-/v/, /w/-/v/, and /ð/-/v/. This study sought to test and expand on the predictions made by Flege's Speech Learning Model (SLM). According to their hypotheses, Italian listeners would perceive English [p]-[v] similar to Italian [p], while Danish listeners would perceive English [v] as phonetically similar to Danish [ʋ][16]. These expectations would result in excellent discrimination for both groups. Furthermore, Italian listeners were expected to discriminate [w]-[v] well since these phonemes are already present in the Italian inventory. On the other hand, Danish listeners were expected to discriminate [w]-[v] less effectively due to their tendency to equivalence-classify both labials as Danish [ʋ]. As for the [ð]-[v] contrast, SLM predicted that Italian listeners would discriminate well because they treat English [ð] as equivalent to Italian [d]. Danish listeners, if they treated these sounds as new L2 categories, were also expected to discriminate [ð]-[v] well. SLM accurately predicted that Italian listeners would perceive all four English target consonants as equivalent to their native categories. However, for Danish listeners, SLM only accurately predicted discrimination for the /b/-/v/ contrast. In conclusion, this study provides valuable insights into how native Italian and Danish listeners assimilate English consonant contrasts perceptually, but SLM shows some flaws when it comes to the Danish perception hypothesis.

[16] voiced labiodental approximant

In a related study, Bohn and Best (2012) investigated how the perception of non-native consonant contrasts in American English is influenced by phonetic and phonological propertiesof the listener's native language. This was studied with native speakers of Danish and German, which have /r 1 j/ but lack /w/. Danish and German speakers realised /j/ identically, but Danish/German 'light' alveolar [l] differs modestly from English 'dark' [] (velarized). Danish listeners' performance on /w/-/r/ and /r/-/l/ approached that of English speakers, with discrimination of /w/-/j/ being higher than English speakers', regardless of their exposure to spoken English. Additionally, the study found that both Flege's Speech Learning Model (SLM) and Best's Perceptual Assimilation Model (PAM) were not completely successful in predicting non-native listeners' discrimination and identification of stimuli from three American English approximant continua, given that they theorised that non-native speech perception is influenced not only by the contrastiveness and phonetic realisations of target phonemes in the listener's language but also by broader factors.

The study conducted by Bohn and Ellegaard (2019) aimed to investigate the efficacy of two measures, namely perceptual assimilation and graded discrimination, in determining the level of perceived similarity between L2 sounds. The study also aimed to explore the impact of L2 experience on these measures and the subsequent identification accuracy of nonnative consonants. Through Experiments 1 and 2, the researchers assessed the perceptual assimilation of English initial fricatives into Danish categories, as well as the participants' perception of the similarity between pairs of English fricatives. Based on SLM hypothesis, both perceived dissimilarity and perceptual assimilation predicted identification accuracy equally well for all the sounds under scrutiny, [f, v, θ, ð, s, z, w, j, r, l, tʃ, dʒ], with one important exception where both measures failed, the one concerning [ʃ] and [ʒ]. It suggested that dissimilarity ratings may be equally effective, if not better, than perceptual assimilation tasks in predicting identification accuracy. Therefore, future studies should consider the use of dissimilarity ratings as a potential substitute for perceptual assimilation tasks, and so did I.

2.2.2 Italian speakers perception and production of English L2

Limited research has been conducted on the perception of English sounds among Italian native speakers of L1. Flege et al.'s study (1999) primarily focused on vowels, while Bohn et al.'s experiment (2011) specifically examined consonants. As this paper is centred on consonant speech perception, I will provide a brief overview of Flege's work, followed by a more detailed analysis of Bohn's research.

Flege et al.'s study (1999) involved an investigation into the accuracy of English vowel production and perception among highly proficient Italian-English bilingual individuals. The assessment of vowel perception was conducted using a categorical discrimination test. A significant finding from this study is that early bilinguals develop new categories for vowels in their second language, and there is a noteworthy correlation between perception and production. This finding supports the hypothesis proposed by the Speech Learning Model (SLM) that the accuracy of L2 vowel production is influenced by the accuracy of their perception. Additionally, the research revealed that the likelihood of category formation for L2 vowels is directly linked to the perceived cross-language phonetic similarity. According to the SLM (Flege, 1995), achieving native-like production of phonetic segments in an L2 is only feasible if they are perceived in a native-like manner. Thus, accurate perception of L2 sounds is a predictor of their accurate production.

Bohn et al. (2011) sought to investigate how L1 Italian speakers perceived three English contrasts: /b/-/v/, /w/-/v/, and /ð/-/v/. They emphasized the disparities in the realization of the three labial consonants /b/, /v/, /w/, which generally exhibit distinct VOT values in both languages. The difference lies in the fact that Italian /b/ is consistently prevoiced as [b], while initial English /b/ is produced as short-lag voiceless plosive that can be perceived by Italians as [p]. Additionally, the Italian phoneme inventory lacks /ð/, which is not similar articulatorily from the Italian voiced dental /d/. It is worth noting that all Italian participants were from Tuscany, which aligns with my own research. They were subjected to an AXB discrimination test, followed by a request to label the stimuli using Italian orthography. The results indicated that English [ð] was predominantly assimilated as the native <d> by 34 out of 41 participants, while three responded with the non-Italian <th>, and four gave varied responses of <d>, <th>, or <l>. Considering that 100% assimilated English [v] as L1 <v>, the discrimination between English [ð] and [v] should be relatively straightforward. From the perspective of Second Language Model (SLM), the study correctly anticipated that Italian listeners would classify all four English consonants (/b, v, w, ð/) as "identical" (equivalent) to native categories. Consequently, they were expected to achieve high discrimination scores for the tested pairs (/b/-/v/, /w/-/v/, and /ð/-/v/). These expectations were met, confirming the SLM represented by the continuum and the theory of "equivalence classification" explained in 2.2. Bohn et al. (2011) also observed that Italian listeners were able to correctly discriminate all the contrasts despite the phonetic differences between English and Italian /b/, as well as between English /ð/ and Italian /d/.

Given the limited number of studies conducted on L1 Italian speakers learning English as L2, my research aims to partially address this gap. Although Bohn et al. (2011) focused on Tuscan speakers, there exist notable differences between their work and mine. Primarily, they employed an AXB discrimination test followed by a labelling task, whereas I will administer two distinct experiments to the participants, namely a graded discrimination task and an identification task. In terms of participants, Bohn et al. (2011) had Tuscan speakers, while I will narrow down the scope to Florentine speakers, which will be further elaborated on in the subsequent explanation. The English sound /ð/, which Bohn et al. analysed, is also central to my research but for different purposes, along with its voiceless English counterpart /θ/.

The next part of this chapter will delve into the specifics regarding the measures that have been described and utilised in Speech Perception, with a particular focus on the types of tests employed to gather data for the purpose of this paper.

2.3 Methods for predicting difficulties

SLM proposes that the perception difficulties experienced by L1 listeners when exposed to nonnative sounds are contingent upon the classification of the L2 sound. According to SLM terminology, this classification can be categorised into three groups: "new," "similar," or "identical." Sounds that fall into the completely new or completely identical categories are more easily identifiable. Furthermore, SLM asserts that production is influenced by perception, implying that the learnability of an L2 sound is linked to how it is perceived by L1 listeners (Bohn & Ellegaard, 2019). In my study, I will discuss the implications of the learnability issue and subsequently make predictions inspired by SLM. These predictions will then be tested through a graded discrimination test (or perceptual similarity test) and an identification test (Chapter 3).

The reason for focusing on difficulties is due to the well-established fact that learning non-native (L2) sounds poses a significant challenge (Bohn, 2018). It is widely documented that this difficulty in learning primarily stems from perception, which Strange (1995, p. 2) refers to as "accented perception." This difficulty is generally understood to arise from various factors related to the learner's L2 background, including exposure, experience, the amount of learning, and individual differences. SLM introduces an additional factor, the relationship between the L1 and L2 phonological systems, which I will refer to as "similarity," especially in the context of the first experiment. The premise is that learnability is predominantly influenced by the perceived characteristics of the sound and its categorization as "new," "identical," or "similar" according to SLM's framework. It is through phonetic similarity that we can anticipate success or difficulties in learning non-native sounds (Rato, 2022).

The topic of L2 acquisition's level of difficulty is discussed by Bohn (2018), who concludes that accurately perceiving and producing nonnative speech sounds can be challenging. Although articulatory factors have traditionally been considered insignificant, alternative explanations have been sought. Flege (1995) emphasises the necessity of empirical measures to assess difficulties and rejects any predetermined conclusions. Bohn (2018) emphasises the paramount importance of valid measures of cross-language similarity, which will be explored subsequently.

In particular, Bohn (2002) delves into the concept of Perceptual Similarity. A variety of experiments have been employed to test this concept. After providing a brief overview of the commonly used methodologies, I will further elaborate on the specific task employed in my study for the experimental phase of the first experiment, which aimed to measure perceptual similarity. However, Chapter 3 will be dedicated entirely to the empirical section, where a comprehensive explanation of the experimental designs and types of experiments will be provided.

2.3.1. Methods used for assessing similarity

Similarity is a concept developed in many fields, including that of phonetics, with Ladefoged (1990) proposing principles for measuring similarity in cross-language research. In contrast to the Contrastive Analysis (CAH) developed in the 1950s-1960s, perceptual similarity is based on the relationship between L1 and L2 in terms of their perceived similarity. Bohn (2002) identifies three types of sounds: stimuli that are identical or sound the same, stimuli that are the same but have different phonetic realisations, and those that are very different and/or have no equivalent. The success in acquiring speech sounds depends on the level of similarity, whether it is "new," "identical," or "similar." It is important to note that this is essentially the foundation of Flege's SLM in 1995. Essentially, the SLM predicts that the ease of establishing a new phonetic category depends on the perceived similarity by the learner when comparing an L1 and an L2 sound. Flege also discusses methods for determining how an L2 sound would be perceived in relation to an L1 sound. In other words, SLM and other approaches use similarity as a fundamental concept for understanding the perception of non-native speech, which is why SLM serves as the reference point for my research.

In the study conducted by Bohn (2002), various methods for operationalising and measuring the concept of "similarity" are exemplified. These methods can be broadly categorised into two approaches: “armchair methods” and “experimental methods”.

Armchair methods primarily rely on comparing the symbols used to represent sounds in different languages. However, these methods are considered to be only preliminary due to their low reliability. This is because the selection of symbols is arbitrary and can vary depending on several factors, including the languages being compared. Additionally, sounds that share the same symbol may differ in terms of quality or quantity, which is not captured by the International Phonetic Alphabet. Nonetheless, Bohn (2002) suggests that armchair methods can serve as a starting point for assessing similarity, with the support of experimental methods.

Experimental methods, on the other hand, encompass acoustic comparison, articulatory comparison, and perceptual assessment of similarity. Among these methods, I will provide a more detailed explanation of perceptual assessment of similarity, as it plays a crucial role in operationalising part of my research question.

Acoustic comparison method and articulatory comparison method

Acoustic comparison and articulatory comparison methods are experimental approaches that involve comparing the acoustic properties and articulatory characteristics of speech sounds, respectively. However, they present practical challenges due to the inherent heterogeneity of sounds, even when variables are controlled. It is nearly impossible to reproduce the exact conditions for each sound, and it has been observed that the same sounds in different languages have distinct properties. Moreover, Bohn (2002) highlights that acoustic comparisons alone cannot answer the question raised by Flege (1995) regarding when an acoustic difference between a learner's first language (L1) and second language (L2) becomes phonetically relevant. In other words, while it is possible to compare sounds based on their physical acoustic properties, this may not reliably predict perceptual similarity. Similar challenges apply to articulatory comparisons, as a physical difference detected in articulation does not necessarily aid in identifying perceptual similarity.

Perceptual assessment of similarity

Perceptual assessment of similarity is the focal point of my research, as it acknowledges that phonetic similarity is relative and influenced by individual factors such as L1, L2 experience, type of stimuli, and task requirements. Despite this subjectivity, extensive research on phonetic similarity has established rigorous procedures for determining the level at which listeners typically compare speech sounds for similarity (Bohn, 2002).

The perceptual assessment of similarity can be approached through either direct or indirect methods. Direct methods involve listeners providing ratings of similarity or dissimilarity for pairs of sounds, while indirect methods do not require listeners to respond directly but instead use indicators believed to be indicative of similarity.

Furthermore, studies on perceptual similarity can be categorised based on the nature of the stimuli. There are explicit tasks, where both stimuli are present, and implicit tasks, where one stimulus is present while the other is stored in memory. These are alternatively referred to as "perceptual similarity" and "ecphoric similarity."

For the purpose of this paper, my focus will be on explicit perceptual similarity tasks, specifically overt and direct tasks. These tasks have been utilised in both intra- and cross-language studies, including the Horslund and Bohn (2022) study, which will be discussed in the following section. In my own research, I will employ a paired comparison method, where participants will be asked to rate the perceived similarity between two sounds. However, it should be noted that triadic comparison methods can also be employed.

While it is important to acknowledge that participants may be influenced by the instructions they receive, it is widely accepted that direct perceptual similarity methods are the most valid. Bohn (2002) affirms that these methods are not only valuable for cross-language studies on speech sound similarity, but also for predicting perceptual issues and second language (L2) speech perception.

2.3.2 Types of test used in my study

Graded discrimination tasks

Discrimination tasks are commonly used to assess the listener's capacity to differentiate between two or more sounds. According to McGuire (2010), discrimination tasks are highly effective for determining an individual's perception with regard to variables such as their first language (L1) and exposure to a second language (L2). Several types of discrimination tasks have been developed, including the within-language graded discrimination task, which I employed for experiment 1.

To date, the within-language graded discrimination task has received limited attention. However, Bohn and Ellegaard (2019) have demonstrated its validity in predicting the identification of L2 sounds by non-native listeners. They conducted a study with two groups of native Danish listeners differing in their experience with the English language. The researchers employed two measures of perceptual similarity: an assimilation task and a within-L2 graded discrimination task. These measures were used to predict identification and determine the most effective method for predicting identification accuracy. Regarding the within-language graded discrimination task, participants were asked to assess the similarity of selected English fricative pairs using a 9-point Likert scale ranging from 0 ("identical") to 8 ("very different").

Apart from the results, Bohn and Ellegaard (2019) not only found that the within-language graded discrimination task effectively predicted identification accuracy but also recommended its increased utilisation in future studies. This is because the task uses ratings rather than potentially ambiguous orthographic labels.

In my study, I implemented a within-language graded discrimination task, also known as the "perceptual similarity task" according to Bohn (2002). This task is specifically designed to reveal perceptual similarity. Participants are required to indicate the similarity between two auditory stimuli using a 5-point Likert scale (ranging from 0 to 4). Following the suggestion of Ellegaard and Bohn (2019), I also employed this task to predict identification accuracy, which was assessed in experiment 2 through an identification task.

Identification task

The second part of my study involves an identification task. In identification tasks, listeners are typically presented with one or more sounds and asked to provide an explicit label for each stimulus. In the basic version of an identification task, a single stimulus is presented during a trial, and the participant assigns a label to that stimulus. The label can be provided by the experimenter, thus requiring the participants to choose from a predetermined set of labels, or the participants may have to independently indicate the label (e.g., "write what you hear") (McGuire, 2010). This type of task is commonly employed to assess categorical knowledge. It is considered to be easy for participants to understand and quick to administer. However, McGuire (2010) highlights the disadvantage of using labels, as it "imposes a categorical decision for the subject." This drawback will be duly addressed in the final discussion and limitations in Chapter 6.

In my particular case, I have devised an identification task aimed at examining the hypotheses posited in experiment 1, as well as making predictions regarding the acquirability aspect. Participants are instructed to associate the auditory stimuli they hear with one of the ten orthographic labels displayed on the screen.

To the best of my knowledge, there have been no prior investigations conducted on the perceptual distinctions in L2 English among L1 Italian listeners, specifically addressing the Florentine dialect for its phonetic variations. Furthermore, I have not come across any other research that employed a graded discrimination task within the same language, followed by an identification task, apart from the aforementioned study by Ellegaard and Bohn (2019).

3. The present study

In Chapter 2, I gave a brief overview of the history of studies in speech perception, followed by an analysis on the contribution of various speech perception studies to the initial stages of research on how Italian speakers perceive and produce English sounds. These were categorised based on whether the focus was on vowel or consonant sounds.

The study conducted by Bohn et al. (2011) focussed on consonant sounds, and it was examined in detail for the following reasons. Firstly, it explores the interaction between specific phonological systems of the native language (L1) and the perception of non-native sounds, specifically English as a second language (L2). Secondly, it considers two different L1 groups, namely Danish and Italian. Although Bohn et al. (2011) examine different languages from the present study, which focuses on dialects within the same country, both his study and my study adopt the same theoretical framework, namely the Speech Learning Model (SLM) by Flege. Thirdly, Bohn's study consists of two experiments, where the first one involves a discrimination task that provides predictions for the second experiment, which is an identification task, specifically a labelling test. Similarly, the present study is structured with two phases, the first being a graded discrimination task that also serves as a predictor for the labelling task in the second phase.

The objective of the present study is to investigate whether graded discrimination within the L2 can effectively predict L2 identification. Additionally, this study explores the potential influence of the L1 phonological inventory on the graded discrimination and identification of L2 sounds. This investigation focuses on the perception of English fricatives /θ/ and /ð/ by native Italian speakers, taking into account their respective dialects. The research methodology includes a within-language graded discrimination test and a subsequent identification test. The tests employed in the present study were introduced in Chapter 2.

Initially, participants were asked to take part in a within-language graded similarity task. The predictions concern the two English sounds of interest, namely /θ/ and /ð/. According to the literature, [θ] and [ð] are identified as allophones of /t/ and /d/ in the Florentine dialect (Bertinetto&Loporcaro, 2005; Calamai, 2017). Therefore, the main pairs of interest for the predictions are /θ-t/ and /ð-d/ in both directions.

It is predicted that native Florentine listeners will perceive these pairs as very similar or identical, compared to non-Tuscan listeners who are expected to perceive a higher level of dissimilarity (different/very different), regardless of the directionality. This prediction is based on the fact that /θ/ and /t/, as well as their voiced counterparts /ð/ and /d/, are allophones in Florentine dialect (Calamai, 2017), and are used in complementary distribution without a change in meaning.

In order to assess the dissimilarity between pairs of sounds, participants in the experiment were instructed to rate them on a scale ranging from 0 to 4, with 0 if they perceive the sounds of the pair as identical, 1 and 2 indicating very similar and similar, and 3 and 4 indicating different and very different respectively. The expectation was that Tuscan participants would rate the pairs as more similar compared to non-Tuscan participants.

The impact of dialect on the contrasts may vary. Specifically, lenition of stops, known as *gorgia toscana*, has a lesser effect on /d/ (Calamai, 2017). This leads to the prediction that Florentine listeners would perceive /ð/ and /d/ as less similar compared to the voiceless counterparts. Consequently, the similarity ratings for all the contrasts of interest would be higher for non-Tuscan listeners, but there would be slight differences in the ratings for voiced and voiceless pairs.

The dissimilarity ratings provide insights into the predicted accuracy of sound identification, which is examined in Experiment 2. For Florentine listeners, the anticipated low dissimilarity ratings for /θ-t/ and /t-θ/ suggest that they may struggle to correctly identify /θ/. Conversely, for non-Tuscan listeners, high dissimilarity ratings indicate that the two sounds in the pair are well distinguished, leading to unambiguous identification by the non-Tuscan group. However, varying degrees of dissimilarity among the pairs can result in different levels of accuracy. In other words, non-Tuscan participants may outperform Florentine participants in identifying both /θ/ and /ð/, but the higher dissimilarity rating predicted for the voiced pair suggests even more accurate identification, with slightly higher accuracy for non-Tuscan listeners.

The second question pertains to the issue of mapping, which relates to the ease or difficulty with which listeners perceive and learn non-native sounds. The Speech Learning Model (SLM) proposed by Flege (1995,2005) and Bohn & Best (2012) incorporates the effects of language-specific phonetic similarities and plays a central role in understanding non-native speech perception. Predictions are made based on the SLM, as detailed in Chapter 2.

According to the SLM, it is possible to predict that a higher dissimilarity rating results in the formation of a new category for the "new" sounds, while a lower rating, indicating similarity between two sounds, leads to a merging of the two categories. In the context of this experiment, this would mean that Florentine participants may more frequently identify /θ/ with /t/ compared to non-Tuscan participants, resulting in the substitution of the sound /θ/ with /t/, and the equivalence of the orthographic representations of <THink> and <t> in the test. The same pattern applies to /ð/ and /d/, though to a lesser extent. In the experiment, Florentine participants may indicate <THem> when hearing /d/ and incorrectly identify the sound /ð/. According to the SLM reasoning (Flege, 1995), Florentine listeners are expected to perform poorly on the identification task because they classify both English /θ/ and /ð/ as phonetically similar to /t/ and /d/, respectively, regardless of the direction of stimulus presentation.

In summary, the following hypotheses have been proposed regarding the sounds of interest, /θ/ and /ð/:

1. Florentine listeners are expected to perceive both contrasts (in both directions) as more similar compared to non-Tuscan listeners. Specifically, the contrast /θ-t/ is anticipated to receive mean ratings between 1 and 2 from the Florentine group, while the non-Tuscan group is expected to rate it between 3 and 4.

2. The influence of dialect is expected to differ depending on voicing. Specifically, the Florentine group is predicted to perceive [ð-d] as less similar compared to [θ-t].

3. Dissimilarity ratings may impact identification accuracy, with lower ratings indicating less accurate identification. It is anticipated that non-Tuscan listeners will outperform Florentines in both the identification of [θ] and [ð]. Additionally, the Florentine group is expected to encounter difficulties in correctly identifying /θ/, whereas the relatively high dissimilarity ratings for [ð-d] suggest fewer issues in correctly identifying /ð/.

4. According to the Speech Learning Model (SLM), non-Tuscan listeners are expected to develop a new category for the "new sounds" /θ/ and /ð/, while Florentines will not. Moreover, it is possible that Florentine listeners may more frequently misidentify /θ/ as /t/ compared to non-Tuscan listeners. The same pattern may apply for /ð/ and /d/, albeit to a lesser extent.

3.1. Experiment 1: within-language graded discrimination

3.1.1. Method

Participants

A total of thirty-eight participants, consisting of both male and female Italian native speakers with a mean age of 30 years old were involved in the research study. Out of these participants, eighteen individuals were from Tuscany and identified themselves as speakers of the Florentine dialect (Florentine group, FI), while the remaining twenty participants were from regions outside Tuscany (non-Tuscan group, NT). The non-Tuscan participants were specifically from areas where dialects other than Florentine are spoken and did not possess any Tuscan accent or dialect. Please note that the data analysis did not include two of the recruited non-Tuscans who actually took part in the experiment. The decision was taken after a thorough examination of all participants' responses concerning the fillers in the first experiment. Specifically, participant number 2 exhibited a preference for assigning a rating of 2 (*similar*) to the majority of the sounds, whereas participant number 28 only used ratings corresponding to 0 and 1. Consequently, they were excluded from the data analysis to avoid the possibility that their overall responses were inconsistent with the task itself.

The recruitment process used social media platforms and personal connections to reach out to potential participants. The average age of the participants was thirty years old, and all of them had acquired Italian as their first language before starting to learn and speak English. It is important to emphasise that their participation in the experiment was completely voluntary and uncompensated.

For the recruitment process, the participants were sent a message containing three separate links to be completed in sequential order. Firstly, they were required to fill out a sign-up form along with an Italian consent form (see Appendix). Following that, they were directed to complete the Language Profile Questionnaire (LPQ), aiming to evaluate their dialectal background and English language usage, before proceeding to the actual experiment (cfr. Appendix for some of the relevant questions). The two experiments were actually presented to the participants in the form of a single experiment divided into two phases. By doing so, it was more straightforward for the participant to follow the guidelines, and there would have been less chances that they did not complete both experiments. If I had sent to them two different links for the two separate experiments, the number of participants not completing one of the two would have been higher, and I would have had less data to analyse. Alternatively, it was less time-consuming for both parts.

The Language Profile Questionnaire (LPQ), conducted through Google Forms, was designed to discern the differences in dialect usage between the two groups and determine the extent of exposure the participants had to spoken English, rather than evaluating their proficiency (and they have been repeatedly made aware of that). The questionnaire included inquiries about their linguistic background, language usage patterns, habits, and opinions regarding English pronunciation. Furthermore, it aimed to gauge their understanding of Italian language phenomena, particularly the concept of spirantisation/lenition (commonly referred to as "aspirazione" among Italians).

The graph below attests that the great majority of the participants (32) still use English daily, whereas 6 of them tend to practise less English. No participant among the 38 individuals whose responses were analysed in the study reported no use of English at all. Language usage was determined based on participants' self-reported frequency (Figure 3).

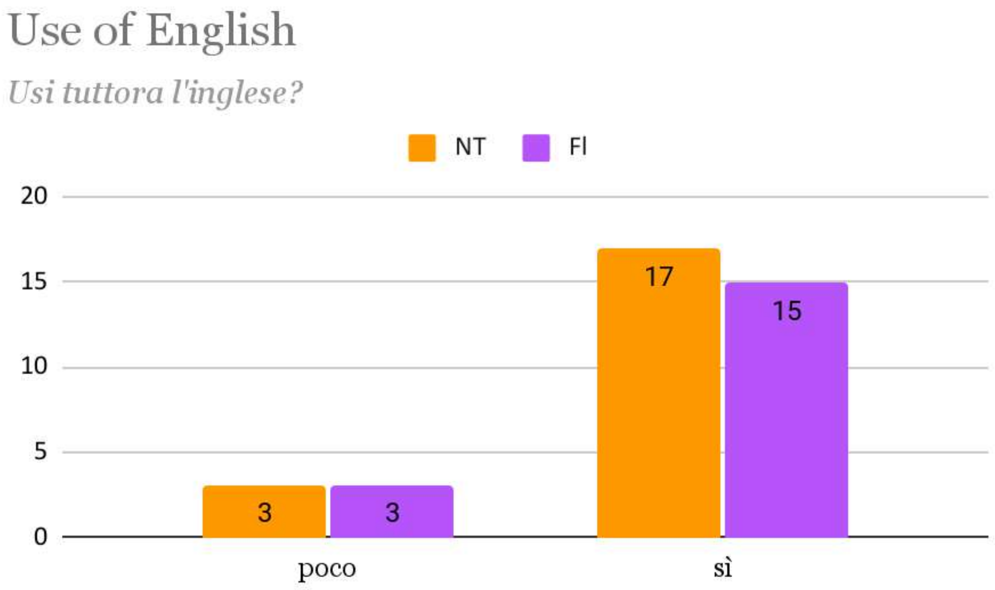

Figure 3

Mean responses about use of English to the question "Usi tuttoral'inglese?" in the LPQ. The possible answers were "sì", "poco", "no".

Less homogeneous is their awareness of or knowledge about the spirantisation of the interdental fricatives[17]. The following graphs show the responses of the participant to the question "Sai cosa è l'aspirazionedella t e/o della d?"[18]. Each group has a separate graph, as to point out the deviations in their answers.

[17]In the Google Moduli LPQ the participant was asked "saicosa è l'aspirazionedella t e/o della d?". I referred to the spirantisation of the interdental fricative as *aspirazione* because it is the commonly known way of referring to it, and I wanted them to understand the question.

[18]The possible answers were "sì, per entrambe", "solo della t", "solo della d", "nessuna delle due".

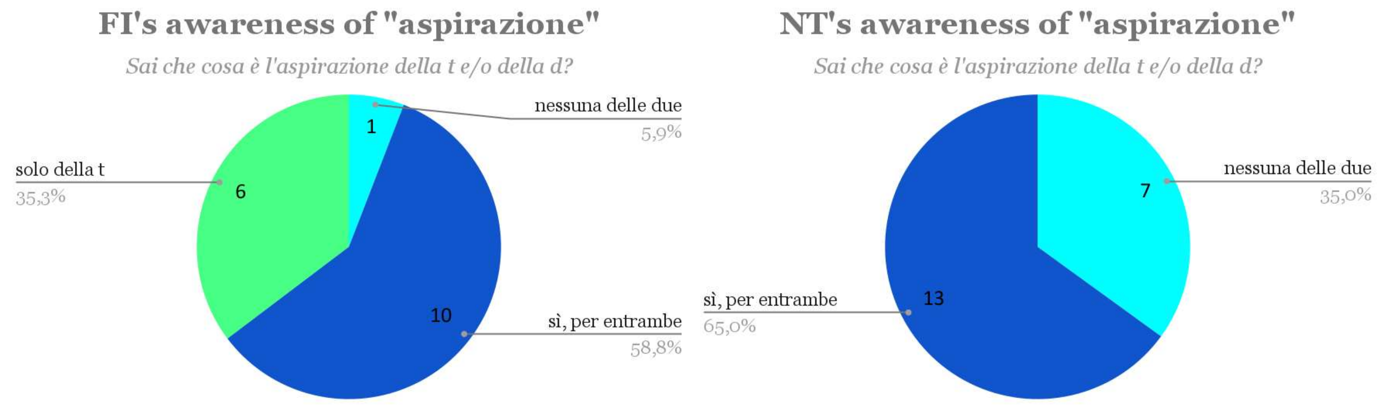

Figure 4

Pie chart showing the participants' awareness of "aspirazione", calculated from the responses to the question "Sai cosa è l'aspirazionedella t e/o della d?" in the LPQ. On the left, Florentine speakers', on the right non-Tuscan ones'.

Twenty-three out of thirty-eight participants know about "aspirazionedella t e/o d". Conversely, seven NT participants do not at all. Six FI listeners, instead, are aware of the phenomenon when it comes to the sound /t/, whereas no individual's awareness was restricted to the spirantisation of the /d/ sound.

In the Appendix section I also provide the graphs taken directly from Google Forms regarding the Language Profile Questionnaire, as for more detailed information about their English language exposure and use in different contexts, and their dialectal profile. Following the completion of the Language Profile Questionnaire (LPQ), participants were directed to *Pavlovia.org* (Peirce et al., 2019) to commence the test. The test on *Pavlovia.org* consisted of two parts, corresponding to Experiment 1 and Experiment 2 in the research: part 1 involved a within-language graded discrimination task (perceived similarity), while part 2 immediately followed part 1 and consisted of an identification task.

Design and Materials

The experiment employed a within-language graded discrimination task (see Chapter 2). The materials consisted of specifically selected English consonant sounds, namely the two English dental fricatives /θ, ð/ under investigation, as well as the English sounds /f, v, s, z, ʃ, ʒ, t, d/. These sounds were presented to participants in [Ca] pairs, such as [f-θ], [v-θ], [θ-ð], [θ-t], [θ-d], [f-ð], [v-ð], [t-ð], [d-ð], each presented also in reverse order, and then twice, for a total of 36 pairs. Additionally, ten type-identical pairs (e.g., [f-f]) were presented twice. The dyads without the target sounds were also included, and they differed only in voicing (i.e. [f-v], [t-d], [s-z], [ʃ-ʒ]. As they are fillers, they appeared twice, but in the same direction of presentation. This resulted in a total of 64 pairs of sounds, with the order of presentation being randomised.

Participants were required to rate the perceived similarity of each randomly presented pair on a 5-point Likert scale ranging from 0 (identical) to 4 (very different). The scale was visually represented by diamonds, which were of equal dimension and equidistant from each other.

Procedure

The experiment was created using the open-source *Psycho.py* software, allowing the upload of the experiment onto *Pavlovia.org* (Peirce et al., 2019) for remote participation by participants. After completing the LPQ, participants accessed *Pavlovia.org* and entered their name to initiate the test. Initially, a welcome screen appeared, followed by instructional slides that provided participants with all the necessary information regarding the test procedure. Subsequently, a practice phase comprising a single trial was presented to familiarise participants with the sound volume and the nature of the task. Participants were then informed that the experiment would commence on the subsequent screen whenever they decided to press the spacebar (as indicated on the screen). The complete instructions for the procedure are reported in Appendix B.

Experiment 1 aimed to examine the graded discrimination of selected English consonant pairs. Listeners were presented with the relevant pairs of interest four times each, along with other pairs presented twice, resulting in a total of 64 pairs of sounds being listened to. The [Ca] tokens in each pair were presented with an inter-stimulus interval (ISI) of 1.0 second. Participants rated the perceived similarity of each randomly presented pair using a 5-point Likert scale ranging from 0 (identical) to 4 (very different). Only by providing a similarity rating using the mouse on the scale could the experiment proceed. Figure 7 provides an illustrative example of the actual experimental procedure. It represents the screen of each participant while they were presenting with a pair of sounds: they had to click on one of the five diamonds according to their perception. The counting of the stimuli was also indicated in the bottom right-hand corner of the screen. When the participant clicked in the diamond, the experiment moved forward automatically.

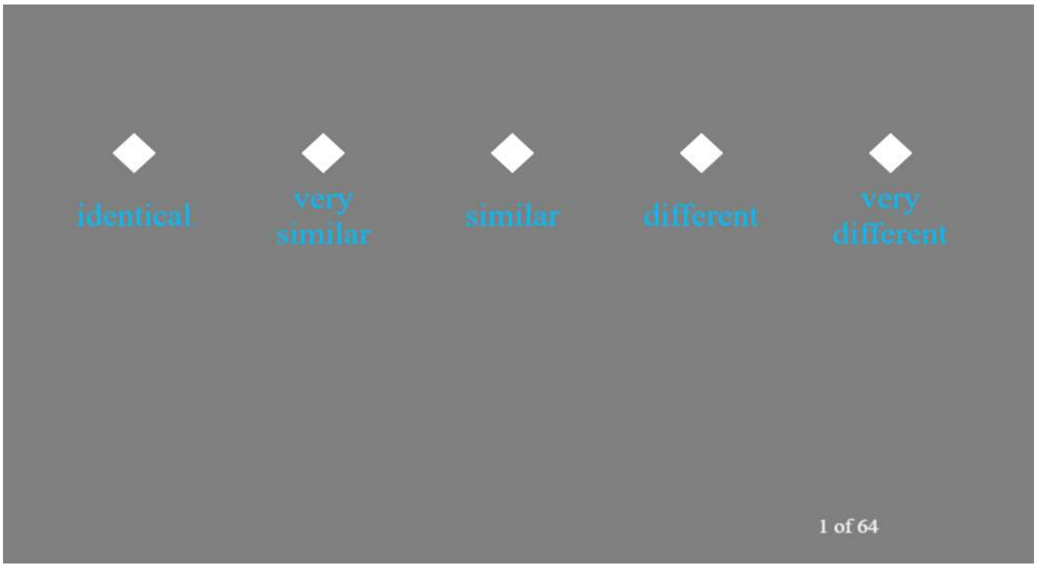

Figure 5

Screenshot of the screen as it was during Experiment 1. The participants heard the sounds, and they always had the scale (diamonds) from identical to very different in front of them. A countdown was also displayed on the bottom-right hand corner.

Data analysis

In the following sections, I will comment on the data collected. The initial step involves conducting a descriptive statistics analysis, followed by an inferential statistics analysis. The data underwent statistical analysis using generalised linear mixed-effects regression models (GLMMs) with the aid of R package *lme4* (Bates et al., 2015).

3.2.2 Results

Filler items

Before presenting the results for the target items, it is pertinent to briefly discuss the responses of both groups to the fillers. The inclusion of filler items in the experimental design served the primary purpose of introducing stimulus diversity and preventing participants from discerning the true nature of the study. Examining the responses to the fillers is instrumental in verifying participants' focus on the task and subsequently including their data in the analysis. Table 4 displays the total number of responses for each filler item, rated on a scale from 0 to 4. For ease of comparison, the pairs with identical sounds are presented first, followed by pairs with distinct sounds.

	0	**1**	**2**	**3**	**4**
d-d	55	16	5	0	0
f-f	59	12	3	1	1
s-s	67	6	1	2	0
ʃ-ʃ	66	8	1	1	0
t-t	55	15	5	0	1
v-v	60	13	1	1	0
z-z	62	9	4	1	0
ʒ-ʒ	70	4	4	1	1
s-z	1	3	9	41	22
ʃ-ʒ	0	4	8	35	29
t-d	13	19	13	24	7
v-f	0	0	9	32	35

Table 4

Total number of responses collected from both FI and NT groups for filler items only.

The results indicate that the majority of dyads with the same sounds were predominantly rated as 0 (identical), with a few exceptions. Dyads with different sounds exhibited a more varied pattern, as they received equal frequency ratings of 3 (different) and 4 (very different), with some instances of a rating of 2 (similar). Notably, the pair /t-d/ displayed a more diverse pattern of choices, which might have to do with the acoustic characteristics of the physical sound.

To sum up, Table 4 shows that most pairs with the same sounds received predominantly a rating of 0 (indicating identical perception), with a few exceptions. The pair /t-d/ exhibited a more varied pattern of choices, possibly due to confusion arising from the difference in voicing[19]. However, this was not observed with pairs such as /s-z/, /ʃ-ʒ/, and /f-v/. In general, the responses to filler stimuli indicate that the participants did not respond indifferently to the stimuli, and both Florentine (FI) and non-Tuscan (NT) participants exhibited highly uniform perception of identical sounds (e.g., pairs like /f-f/), which was the reason why I could continue further with the experiments.

Target items

Before proceeding with the statistical inferential analysis, the mean dissimilarity ratings for the contrasts of interest are provided in Table 5 subsequently (following Bohn & Ellegaard, 2019).

[19] Voicing is apparently the only distinguishing feature. Nonetheless, it is well-attested that VOT is completely different in Italian and English (i.e. Marotta, 2008). Even though the paper is not focussing on VOT, it is important to mention it.

Target[20]	Similarity rating
[ð-d]	2,1
	3,1
[ð-v]	2,7
	3,1
[θ-t]	2,8
	3,6
[θ-f]	1,0
	0,8

Table 5

Mean dissimilarity ratings for the target sounds in the pairs of interest. Violet indicates FI, whereas orange NT.

The Florentine listeners (violet) rated all four contrasts as more similar than the non-Tuscan listeners (orange), except for the [θ-f] pair. However, the effect of dialect varies across the four contrasts. Table 5 suggests that the Florentine inventory may impede the perception of dissimilarity most for [θ-f], to a lesser extent for [ð-d], [ð-v], and [θ-t]. The two groups differ greatly in their ratings for both [ð-d] and [θ-t], which would suggest a group effect. The mean rating for [ð-v] do not differ much in the two groups, and definitely neither as for [θ-f]. To be noted is that this is the only case where the mean rating is lower in NT's responses than FI's.

To test the effect of dialectal variety on the perceived similarity of English interdental fricatives, a linear mixed-effect regression model was built in R (R Core Team, 2018) using

[20]Order of presentation is not relevant for the purpose of the calculated rating means.

lme4 (Bates et al., 2015). To analyse the responses to the target stimuli elicited from 38 participants, a GLMM[21] was fitted with the similarity rating as the dependent variable; random intercepts for participants; directionality (whether the interdental sound was presented firstly or secondly); voicing of the sound (voiced vs. voiceless); place of articulation (abbreviated as poa, i.e. dental, interdental); group (Florentine, FI vs. non-Tuscan, NT); and the interaction between voicing and poa as fixed effects. The descriptive statistics and regression coefficients are summarised in Table 6. The random intercepts is the number of participants, corresponding to 38. The number of observations is 608 and it is restricted to the target stimuli, which is 16 per participant.

[21] Generalised Linear Mixed effect Model

	Rating		
Predictors	*Estimates*	*CI*	*p*
(Intercept)	2.63	2.39 – 2.88	<0.001
Directionality [interdental presented second]	-0.15	-0.28 – -0.01	0.033
Voicing [target sound voiceless]	0.49	0.29 – 0.68	<0.001
Place of articulation [non-target sound labiodental][22]	0.13	-0.06 – 0.32	0.180
group [NT]	0.40	0.12 – 0.68	0.005
th_dh [voiceless] * poa [labiodental][23]	-2.52	-2.79 – -2.25	<0.001
Random Effects			
σ^2	0.73		
$\tau_{00participant}$	0.15		
ICC	0.17		
N participant	38		
Observations	608		
Marginal R^2 / Conditional R^2	0.509 / 0.591		

[22] Place of articulation abb. *poa*
[23] th-dh [target sound voiceless] * place of articulation [non-target sound labiodental]

Table 6

Regression results for responses to target items in within-language similarity rating experiment (table was created by using Sjplot package).

A robust effect of *Voicing* was attested in the data ($p < .001$). When comparing this result to the descriptive statistics table of similarity ratings (Table 5), it is evident that the ratings for voiceless dyads, when combined, are lower than those for voiced dyads overall. This can be attributed to the significantly low mean rating for the dyad /f-θ/. The regression table data supports this finding, as it indicates that the voiceless sound (/θ/) is more commonly perceived as similar, highlighting a robust effect of voicing.

The interaction between voicing and poa is also significant ($p < .001$), meaning that when the sound is voiceless and labiodental is perceived more similar than other combinations (strong effect of voicing and poa). Furthermore, Fl participants rated the target sounds more similar than NT participants did ($p = .005$). In addition, the order of presentation of the stimuli affected similarity ratings to a certain extent. When the interdental was second in the stimulus pair, the two sounds were considered more similar ($p = .033$). Conversely, no significant effect of place of articulation alone was found ($p = .180$).

Based on these findings, it can be hypothesised that NT individuals should perform better than Fl individuals for [θ]. Both groups, however, should easily identify /θ/ when in contrast to /t/, which contradicts the initial hypothesis 1. Additionally, the low ratings from both groups for /f-θ/ indicate reduced accuracy and potential difficulties in identifying the target sound. On the other hand, the high dissimilarity ratings for /ð-v/ by both Fl and NT suggest relatively accurate identification for both groups, with slightly higher accuracy expected for NT.

The most significant differences between the two groups are observed in their ratings for /t-θ/, with lower dissimilarity ratings from FI predicting lower accuracy in identifying [θ], and higher ratings from NT predicting more accurate identification. The very similar and relatively low dissimilarity ratings for /f-θ/ from both groups suggest that both Fl and NT individuals will encounter challenges in correctly identifying /θ/, with the possibility of perceiving it as /f/.

The statistical analysis suggests a strong effect of voicing and interaction between voicing and poa, and, in line with hypothesis 1, a robust effect of group, meaning that the two groups show differences in the perception of the sounds of interest. This finding would predict differences in the identification of the tested sounds as well, with lower ratings indicating less accurate identification in Experiment 2.

Hypothesis 3 anticipated that non-Tuscan listeners will outperform Florentines in both the identification of /θ/ and /ð/, and I still expect the same result based on the findings of Experiment 1. However, hypothesis 4 is already partially disproven. I had predicted that /ð/ and /d/ would be misidentified more frequently by FI than NT, which still holds, but to a lesser extent than /θ/ and /t/. These results have led us to some of the observations connected to the research questions, which will be further discussed in Chapter 4. Furthermore, they have been used to make predictions about identification patterns and accuracy, to be tested in Experiment 2.

3.2. Experiment 2: within-language identification

Experiment 2 examined Florentine and non-Tuscan speakers' identification of the same 10 English consonant sounds as in Experiment 1.

Predictions derived from the perceived similarity data

Perceived similarity rating results yield prediction for identification accuracy of the same sounds, as was proved by Bohn and Ellegaard (2019). Based on the data collected from Experiment 1 concerning only target sounds, we can make reference to Table 5, where the similarity ratings of the following target sounds are displayed: [ð-d], [ð-v], [θ-t], [θ-f].

The similarity ratings yield the following predictions for identification accuracy. NT should outperform FI for [θ], but both should differentiate it easily from [t], which is contrary to our hypothesis. At the same time, the low ratings by both groups for [f-θ] predict reduced accuracy and difficulties in the identification of the target sound. The high dissimilarity ratings for [ð-v] by both FI and NT predict fairly accurate identification for both groups, with slightly higher accuracy for NT than FI. The two groups differ the most in their ratings for [t-θ], with lower dissimilarity ratings by FI predicting low identification accuracy for [th], and higher ratings predicting more accurate identification by NT. The very similar and quite low dissimilarity ratings for [f-θ] by both groups suggest that both FI and NT will have problems identifying [th] correctly, and predicts that it might be identified as /f/. The predictions will be checked using error matrices combined with statistical analysis of the identification results in the following sections (Table 7 and Table 8).

3.2.1 Method

Participants

The same participants who took part in Experiment 1 were also involved in Experiment 2. A total of thirty-eight individuals participated. Eighteen of them were native speakers of Florentine dialect with English as their second language, and they resided in and around

Florence. Additionally, there were twenty non-Tuscan speakers of English who participated in the study. The exclusion of two participants from the data analysis was necessary due to specific reasons, already presented in 3.1.1.

Table 3, which has been provided earlier in Experiment 1, assesses the total use of English by the participants. Comprehensive details regarding their L2 experience and dialect use are presented in Appendix A. The recruitment of participants was carried out through various channels, including social media and personal contacts. The mean age of the participants was recorded, and it was observed that all of them had started learning and speaking English after acquiring their first language, which was Italian in all cases. The participants willingly volunteered to be a part of the experiment and did not receive any form of compensation.

Design and Materials

Experiment 2 aimed to investigate the accuracy of native Florentine speakers compared to non-Tuscan speakers in identifying the same 10 English initial consonants that formed the pairs of sounds in Experiment 1. These consonants included /f, v, θ, ð, s, z, ʃ, ʒ, t, d/. However, in Experiment 2, the consonants were presented individually, one at a time. Each sound, except for [s, z, ʃ, ʒ], was presented six times, while the latter set was presented three times as they served as fillers. In total, participants were exposed to 48 stimuli.

Additionally, the screen displayed 10 orthographic English response alternatives corresponding to the audio stimuli. It is important to note that for some consonants, specifically [θ, ð, ʒ], keywords were provided to disambiguate the alternatives. For example, the keyword <THem> was used for [ð], <THink> for [θ], and <beiGE> for /ʒ/, adding to *f, v, s, z, sh, t, d,* and Figure 8 below shows an example of the screen of the participants. This approach was adopted due to the lack of clear and unambiguous orthographic symbols for these particular consonant sounds in English, as was suggested in the study by Horslund & Bohn (2022). The complete instructions for the procedure are reported in Appendix B.

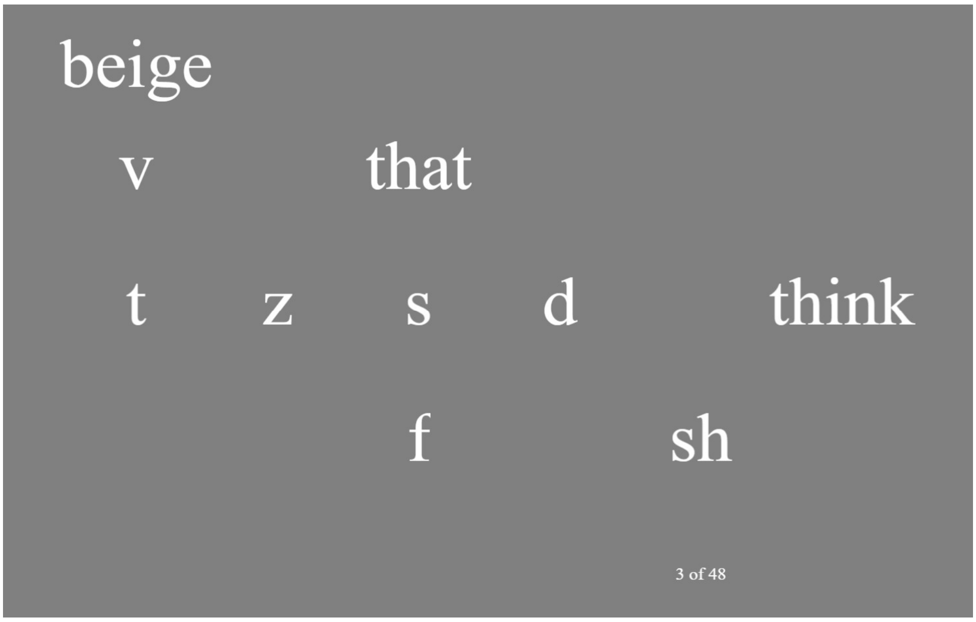

Figure 6

Screenshot of the screen as it was during Experiment 2. The participants heard the sounds, and they always had the orthographic responses in front of them, always in the same position and same colour. A countdown was also displayed on the bottom-right hand corner.

Procedure

Experiment 2 followed the completion of Experiment 1. Participants were given the option to proceed to the second part of the experiment, Experiment 2, by pressing the spacebar. The screen presented them with the 10 orthographic English response alternatives, with some alternatives being provided as keywords to aid in disambiguation. Participants were then familiarised with the stimuli and subsequently listened to recorded sounds. They were required to indicate the sound they heard that corresponded to the correct orthographic response on the screen, based on their perception.

During the experiment, participants listened to English [Ca] syllables and were tasked with identifying the English consonant sound they believed they had heard. Each sound, except for [s, z, ʃ, ʒ], was presented six times, while the latter set was presented three times as they served as fillers. In total, participants were exposed to 48 stimuli. The available response options consisted of 10 orthographically presented consonants, with certain alternatives being provided as keywords to aid in disambiguation due to the absence of an unambiguous orthography for [θ, ð, ʒ] in English. Listeners responded by clicking buttons on the computer screen.

3.2.2 Results

In order to analyse the identification task results, particular emphasis was placed on the determination of correct or incorrect responses. A total of 1864 stimuli and corresponding participant responses were examined. The findings revealed that out of these, 1315 stimuli were accurately identified, while 509 were errors, irrespective of group.

Subsequently, I directed my attention towards the latters. This led me to investigate potential variations arising from dialectal groups. I classified the errors differentiating between those attributed to the Fl group and those associated with the NT group. These categorizations were then presented in the error matrices outlined in Table 7 and Table 8. It can be observed that the distribution of errors is similar in the two groups. Moreover, it is clear that the target stimuli /θ/ and /ð/ are subjected to error more frequently than other sounds, except for /t/ and /d/. A deeper analysis on the target stimuli is needed.

Stimuli	**d**	**ð**	**f**	**s**	**ʃ**	**t**	**θ**	**v**	**z**	**ʒ**	**tot**	**%**
d		40				7	5			17	69	27,5
ð	23					1	4			1	29	11,6
f		2					4				6	2,4
S					1						1	0,4
ʃ				4							4	1,6
T	9	19		1			1			1	40	15,9
Θ		28	60			1					89	35,5
V		3									3	1,2
Z		4		3							7	2,8
ʒ									3		3	1,2
TOT											**251**	100

Table 7

Error matrix for Florentine speakers. Note: the numbers indicate the total instances of errors for each stimulus. The stimuli are in the first column; the responses are in the first row.

Stimuli	**d**	**ð**	**f**	**s**	**ʃ**	**t**	**θ**	**v**	**z**	**ʒ**	**tot**	**%**
d		29		1	2	29	3			18	82	31,8
ð	20						3				23	8,91
f		2					5				7	2,71
s					2					1	3	1,16
ʃ				3						2	5	1,93
t	3	16		1			9				29	11,2
θ		12	74		2						88	34,2
v		5					3				8	3,10
z				3						1	4	1,55
ʒ				1	8						9	3,49
TOT											**258**	100

Table 8

Error matrix for non-Tuscan speakers. Note: the numbers indicate the total instances of errors for each stimulus. The stimuli are in the first column; the responses are in the first row.

The /θ/ stimulus is the most frequently misidentified sound in both groups, with 89 misidentifications for the Fl group and 88 for the NT group. The primary substitute sounds identified by both participants are /f/, /ð/, and /t/ in the case of the Fl group, and /ʃ/ is identified twice as often in the NT observation.

The /ð/ sound is correctly identified more frequently by both groups, with 29 correct identifications for the Fl group and 23 for the NT group. However, it is often confused with the sound /d/. Additionally, the Fl group identifies /d/, /t/, /θ/, and /ʒ/ as substitutes, while the NT group identifies /θ/ only three times.

In addition to the target sounds, it is evident that both /t/ and /d/ show a high percentage of misidentification. Interestingly, this pattern of error is the opposite of their fricative counterparts. Specifically, the voiceless /t/ is more frequently identified correctly compared to the voiced /d/. Surprisingly, the voiced /d/ is an unexpected source of error, with almost as many misidentifications as the /θ/ (69 for the Fl group and 82 for the NT group). It is worth noting that the identified substitutes for the stops /t/ and /d/ are not as clear-cut as other analysed sounds, as they are frequently substituted with /d/, with /ð/, /θ/, and /s/ by both groups, and additionally with /zh/ by one participant from the Fl group. Also, /t/ and /d/ were the two fillers that, when compared, showed a varied pattern of dissimilarity ratings (cfr. Table 2). Conversely, the /d/ sound is misidentified with /ð/, /t/, /θ/, /ʒ/, /s/, and /ʃ/ by both the Fl and NT groups.

	Errors				
Stimulus	*Fl*	*%*	*NT*	*%*	*tot*
ð	29	11,6	23	8,91	52
θ	89	35,5	88	34,2	177

Table 9

Number of errors and related percentage for the target items in the identification experiment indicating both groups.

To analyse the responses to the target stimuli elicited from 38 participants, a GLMM was fitted with the correctness of the response as the dependent variable; random intercepts for participants; voicing of stimulus (voiced vs. voiceless/unvoiced); group (Florentine, Fl vs. non-Tuscan, NT) as fixed categorical effects. The data of interest were restricted to the target stimuli, thus I left out all the stimuli except for /θ/ and /ð/. The descriptive statistics and regression coefficients are summarised in Table 10. The number of observations is 456.

		Correct	
Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	6.33	3.14 – 12.73	<0.001
stimulus_voiced [correct]	0.05	0.03 – 0.08	<0.001
group [NT]	0.55	0.23 – 1.30	0.175
Random Effects			
σ2	3.29		
τ00 participant	1.23		
ICC	0.27		
N participant	38		
Observations	456		
Marginal R2 / Conditional R2	0.350 / 0.526		

Table 10

Regression results for responses to target items in the identification experiment (table was created by using Sjplot package).

According to the data, when the sound is voiced (/ð/) it tends to be more often correctly identified ($p < .001$), thus a robust effect of voicing was found. On the other hand, there is no significant effect of group ($p = .175$), as both groups showed a similar distribution of error.

The inferential analysis combined with the error matrices of Table 7 and 8 illustrate that the Florentine speaking group showed less accurate identification than non-Tuscan listeners for the two target fricatives, but with a small difference that was not significant according to the statistical analysis. An interesting and unexpected exception is that of /d/, because it was incorrectly identified by both groups a number of times that was comparable to the fricative /θ/. I suggest that there might be a connection to what I anticipated in Chapter 1, namely to the fact that English and Florentine languages share certain plosive sounds, but the actual place of articulation and the Voice Onset Time for the /t/ and /d/ sounds is different. In English, these sounds are realised as alveolar sounds, while in Florentine (similar to Standard Italian), they are pronounced closer to dental sounds. Some scholars suggest that this difference may lead to mispronunciation of English sounds (Wheelock, 2016).

The findings of Experiment 1 yielded an unexpected outcome regarding the ability of Florentine speakers to correctly identify [ð], while indicating that the identification of [θ] was comparatively more problematic, often mislabelled as /f/. Consistent with expectations, NT participants demonstrated better accuracy in identifying [θ] compared to Fl participants, although overall accuracy was still reduced, aligning with the low dissimilarity ratings. Both listener groups demonstrated relatively high accuracy in identifying [ð], which was not well predicted by the perceived similarity test. The impact of dialectal differences on the identification accuracy of these fricatives was not observed, as both groups displayed very similar patterns of identification errors. Contrary to the expectations generated by Experiment 1, there was no group effect ($p=0.175$), meaning that the Florentine inventory did not have influence on a different identification of /θ, ð/, even though they are present in the Florentine phonological inventory.

4.General results and Discussion

The study examined in the first instance the influence that a different phonological inventory of Florentine learners of English as L2 could have on the perception of the English sounds /θ/ and /ð/. In general the two groups showed different similarity gradings for the target sounds, confirming the research question that there is a difference in perception between Florentine and non-Tuscan listeners, which might be attributed to the influence of their phonological inventories.

Then, it tested the possibility of predicting identification accuracy of the target sounds from within-language perceived similarity patterns. Due to group differences in the perception, I expected a different pattern of error in the identification patterns of the target sounds between FIs and NTs, but this was disconfirmed. To further elucidate these findings, the four hypotheses presented in Chapter 3 are tested.

The findings of this study provided only partial support for H1. It was expected that Fl individuals would perceive both contrasts (in both directions) as more similar compared to NTs. Specifically, the contrast between /θ-t/ was thought to receive average ratings ranging from 1 to 2 from Fls, while the NTs were expected to rate it between 3 and 4. In terms of the average ratings given by the FIs, it was observed that the gradings for the voiceless interdental fricative fell between 2 and 3, rather than the hypothesised range of 1 to 2. However, it is noteworthy that the non-Tuscan group exhibited higher mean ratings, as hypothesised, which confirms the presence of a group effect for perceptual similarity.

On the other hand, H2 was completely refuted by the results. It was initially hypothesised that the influence of dialect would vary depending on the voicing, suggesting that the Florentine group would perceive /ð-d/ as less similar compared to /θ-t/. However, contrary to expectations, the outcomes demonstrated an opposite pattern of perception for Florentine and also for non-Tuscan speakers. The perceptual similarity mean ratings were lower for the voiced pairs /ð-d/ and /ð-v/ compared to the voiceless dyad /θ-t/, but higher than /θ-f/. The regression model confirmed a group effect and reported a significant effect of voicing, with voiceless perceived more similar than voiced, which is correct only if we consider the voiceless dyads together (/θ-t/ and /θ-f/). I had not predicted such a robust effect of labiodental fricative, but given the very low ratings, I would then expect a high percentage of incorrect identification of /θ/.

It is well attested that the pairs /θ/-/f/ and /ð/-/d/ are some of the most difficult contrast pairs to distinguish (Miller & Nicely, 1955; Reis et al., 2008; Paradis &Lacharité, 2012). This is so common that also L1 English adult speakers, especially under noisy conditions, confuse the fricatives /θ/ and /f/ (Reis et al., 2008; Paradis &Lacharité, 2012).

Acoustically, the two sounds show similar properties. In Appendix D I inserted the three spectrograms from the actual stimuli /θ/, /f/, /t/ used in the experiments. If compared, it can be easily detected such similarity, not necessarily by phonetitians or linguists. In more technical terms, /θ/ and /f/ are non-sibilant fricatives and are different from sibilant fricatives (/s/ and /z/) in that the former exhibit a more or less flat spectrum and have very low intensities. When examining the sound spectrum of fricatives, we find that there are greater resemblances between labiodentals and interdentals than there are between interdentals and coronal sibilants (Miller & Nicely, 1955; Paradis&Lacharité, 2012). These two characteristics combined together make it very difficult to distinguish /θ/ and /f/.

H3 was partially refuted. Part of the prediction that low ratings suggested less accurate identification was true for /θ/ for both groups, in that they were predictors of the misidentification of the sound /θ/. However, I had also put forward that NTs would demonstrate superior performance compared to Florentines in correctly identifying the target sounds, but they resulted in similar identification error patterns, and no group effect. In other words, lower ratings did not result in less accurate identification compared to higher ratings.

Moreover, it was expected that the Florentine group would face challenges in accurately identifying the sound /θ/, while fewer difficulties in correctly identifying /ð/. Actually, there was no significant difference in the identification of the voiced and voiceless target sounds, and neither between groups. This is contrary to expectation based on graded similarity ratings and regression analysis.

What is observed is a general trend of misidentification of [θ] and [ð], irrespective of group. This might be tied to universal acoustic variables that are responsible for the misperception of /θ, ð/, and consequently misidentification of the two (Brannen, 2002, citing the Hancin-Bhatt's model). Even though the adaptation of /θ, ð/ is reported to vary between dialects of a language, it is more common to observe a dominant substitution within a single language irrespective of the dialect (Paradis &Lacharité, 2012), as can be attested in my study where Florentine and non-Tuscan exhibited a comparable identification pattern of errors, in spite of different dialect and different registered perceived similarity.

Hancin-Bhatt (cited by Brannen, 2002) proposed an identification test to different groups of speakers, including native English speakers, where the target sounds were [f, v, θ, ð, t, d, s, z]. The findings of this study indicated that, in terms of onset position, the English group properly perceived /θ/ 64% of the time, followed by the Japanese group (33%), the German group (47%), the Hindi group (48%), and the Turkish group (36%). This highlights that native English speakers themselves perceive the interdental fricatives far from always correctly. This leads me to the conclusion that the pattern registered in my study is ordinary on a universal scale. The same groups tested reported to confuse the interdentals with labiodental fricatives (Brannen, 2002), the same that I observed in Experiment 2.

Specifically, I expected that the Florentine group would experience difficulties in correctly identifying /θ/, but it was interesting to observe a similar pattern for the non-Tuscan group as well. In regard to [ð-d], it was hypothesised that fewer issues would arise in correctly identifying /ð/, and the results confirmed this hypothesis. However, no significant differences were observed between the two groups. Again, this has been attributed to universal trends (Brannen, 2002; Reis et al., 2008).

Moreover, both groups most frequently identified the interdental fricative voiceless as the dental fricative voiceless /f/, but the interdental fricative voiced never as the dental fricative voiced /v/, but rather as /d/. The latter was part of the initial hypothesis, the former is an unexpected result. The literature predicts substitution with /t, d/ respectively, due to a normal phonological adaptation of interdental fricatives to the plosives /t, d/ (Paradis&Lacharité, 2012).

Regarding the phoneme /θ/, Brannen (2002) conducted a study on the phenomenon of "differential substitution". The study aimed to test the hypothesis that substitution of the interdental fricative involves both contrastive and non-contrastive features. She suggested that listeners' choices in attending to specific features during speech perception are influenced by their native language's phonemic inventory. These features are then modified and applied to another sound that is familiar to the listener, a process referred to as "differential substitution". It is important to note that the reason for substitution lies not in the phonemic inventory itself, but in the combination of non-contrastive phonetic representation and contrastive phonemic features. Additionally, the relative prominence of a particular characteristic can be influenced by another characteristic that co-occurs with it. In simpler terms, one characteristic can affect the perceived significance of another characteristic, either diminishing or attenuating it. This mechanism chooses the native representation that is most similar to the target segment based on an algorithmic assessment of how close they sound. This can be applied to the results of experiment 1 and 2, given the high similarity attributed to /θ/ and /f/, and the subsequent misidentification of the target sound. The substitution of the target sound /θ/ with a labiodental fricative might be a result of perceptual confusion involving the features that are considered salient, even though non-contrastive, such as "strident" and "mellow", as Brannen (2002) suggested in her study. Being intensity cues more salient than place cues, it is explained why /θ/ is often identified with /f/.

It has to be noted that there is less research on voiced fricatives than on voiceless fricatives. Voiced fricatives differ from voiceless fricatives in that, in addition to the noise source located in the fricative cord, there is often also a periodic source located in the glottis. The spectrum of non-audible sounds is quite flat as in the case of voiceless sounds (Miller & Nicely, 1955).

One potential factor that may influence the pronunciation of the voiced interdental sounds has to do with morphology. The way interdental fricatives are perceived can be affected by the manner in which they are cognitively assessed and by whether they appear in content words or function words (Reis et al., 2008). The sound /ð/ tends to occur predominantly at the start of words that are classified as function words, as in Experiment 2, where the written option to choose from in the screen was <THat>. Function words, being brief and often less emphasised in continuous speech, can sometimes be mispronounced without significantly impacting comprehension, as their meaning can still be understood from context or may go unnoticed. Listeners do not pay much attention to the correct form, resulting in incorrect identification unintentionally.

Another factor that may account for a uniform pattern of errors, shared by the two groups, has to do with the teaching method they received. What can make a difference is also the accuracy of the phonological input an individual receives at school or by the learning environment (Flege et al., 1999). The participants in my study have not been asked about their former education in detail, but I guess that they had received similar education, given that all of them have studied in Italy, where the teaching system is fairly homogeneous.

The results do not support H4, and therefore the thesis of the SLM is not fulfilled. One of the predictions made by the SLM was that L2 learners are more likely to develop a phonetic category for sounds that are perceptually distant from the nearest native category, assuming all other factors are equal. This cannot be confirmed by the present research, as we found no group effect. This is probably due to lack of experience, which is intended as extensive exposure to English, usually meaning that the listener has been living in an English-speaking country (Guion et al., 2000). The role of familiarity has an impact on perception and learnability. In essence, the extent to which the listener has been previously exposed to the correct target sounds can contribute to their level of familiarity (Polka, 1991). The design of my study actually wanted to make sure that none of the participants had ever lived in an English-speaking Country for an extended period of time, so as to control variability. Being the participants equal from this perspective, lack of exposure as intended above might be one reason why they reported similar identification errors and no group effect.

As for the learning issue, The SLM posits that second language learners need to discern phonetic variations between the sounds of their native language (L1) and the target language (L2) before they can start developing a distinct category for the new L2 sounds. Since the participants of my study reported the target sounds /θ/ and /ð/ to be similar to /f/, /t/, /d/, that is why they did not start developing a new category for the target L2 sound. Again,the development of these categories may require a substantial amount of exposure to native speakers over an extended period of time, particularly for adult learners (Guion et al., 2000).

It is probable that the perception of target sounds in this specific study is influenced by phonological inventories to a limited extent, specifically for /θ-t/, and that the above mentioned factors play an even stronger effect when it comes to perception of other sounds and general identification. Polka (1991) mentions the acoustic features that differentiate sounds as crucial for perception and identification, which was already discussed with Brannen (2002).

The findings of a study conducted by Werker (1995) support the absence of group effect in the identification experiment. The study examined the discrimination by English speakers of two Hindi contrasts. One was the voicing distinction between breathy voiced versus voiceless aspirated stops, /d^h, t^h/; the other one was the retroflex versus dental voiceless unaspirated stop contrast /t̪, ʈ/. While these sounds do not exist in English, English speakers may have some level of exposure to the former sounds in an allophonic context. On the other hand, the contrast between retroflex and dental voiceless unaspirated stops /t̪, ʈ/ of Hindi is less likely to have been encountered by English speakers in an allophonic manner. Nonetheless, the allophonic experience did not lead to better identification. The potential explanation for the limited reflection of phonetic experience in the perception of the Hindi contrasts is probably the fact that phonetic factors may not be relevant in accurately characterising experiential effects, or that they are not the sole factors that need to be taken into consideration.

In general, the discrepancies observed in the perception of the four Hindi contrasts by English listeners were not anticipated when examining the phonemic status, phonetic experience, or acoustic salience factors independently. While the potential significance of phonemic status and phonetic experience has not been disregarded, out of the three individual factors examined, the initial prediction based on acoustic salience was the most consistently aligned with the observed differences in perception, albeit not entirely perfect.

Likewise, the fact that /θ/ and /ð/ are part of the allophonic inventory of Florentine Italian does not directly guarantee perception at ceiling level and consequently identification and learning. On the contrary, given that the results show poor perception and incorrect identification, and that /θ/ was misidentified with /f/ irrespective of dialect, the prediction, relying on the notion of acoustic salience, displayed the highest degree of consistency when compared to the observed variations in perception.

5. Conclusions, Limitations and Future directions

This study represents a pioneering attempt to explore the perception of English sounds by Florentine speaking participants. The aim was to determine whether there were any differences between Florentine participants and other individuals who speak Standard Italian, a group who do not share the same phonological inventories. Specifically, I wanted to answer the main question as to whether the fact that Florentine speakers use [θ] and [ð] as allophones of /t/ and /d/ could influence the perception of those sounds and subsequently alter identification patterns.

As hypothesised, a perceived similarity test conducted within the Florentine group revealed that they perceived the target sounds to be more similar than the non-Tuscans did. This effect was particularly pronounced for the voiceless interdental fricative, though this finding was tentative. This study employed the results of the similarity test to examine identification accuracy and subsequent difficulty patterns, following the approach of Bohn and Ellegaard (2019). However, unlike their findings, the perceived similarity test results in this study did not serve as a reliable predictor of identification accuracy. The identification test demonstrated no discernible group effect, indicating that a similar pattern of identification errors emerged regardless of the participants' group inventory.

It should be noted that the influence of the phonological inventory, which appeared to affect perceived similarity, did not have any impact on the identification of the same sounds. Moreover, the error matrices revealed an unexpected tendency to misidentify the sound /θ/ as /f/, rather than the anticipated misidentification as /t/. This finding confirms that the identification of certain sounds is not determined by the phonological inventory, specifically the use of [θ] and [ð] as allophones of /t/ and /d/ by Florentines. Instead, it appears to be more dependent on the general Standard Italian inventory, exposure to English, and acoustic properties.

Based on the findings from the LPQ survey, it is evident that participants exhibit a shared English language exposure, with all of them commencing their English studies in Italian schools. The similarity in teaching methods employed could potentially account for the prevalent error of misidentifying /θ/ sound as /f/ and vice versa. My data indicates that this misidentification is not influenced by the phonological inventory, but rather by the phonemic inventory, as well as language exposure and teaching methods. The acoustics of sounds are relevant as well, and might play a role in the identification accuracy (or lack thereof). To validate this assumption, further investigation is recommended through comprehensive inquiries into participants' backgrounds, and experimental analysis of sound perception and identification in a range of contexts.

The experimental design of the present study is limited in several ways. Most importantly, by focusing on isolated sounds, in the [Ca] form, the work does not provide a comprehensive picture of the real perception of the sounds as heard in ordinary situations. The importance of context has been emphasised in phonetics literature as for example the influence that neighbouring sounds have between each other, and consequently how this can change the perception (phonotactics studies, i.e. Bohn &Steinlen, 2003). Furthermore, the number of stimuli and the resulting data collected may not be extensive, but considering the experiment's scope, available time, and resources, it is not inadequate either. Some criticism can also be directed towards the orthographic responses displayed on the screen in experiment 2, particularly for the sounds associated with words such as *think*, *that*, and *beige*. The inspiration for using these words stems from the study conducted by Horslund and Bohn in

2022 mentioned in the paper, where similar words were employed, albeit with *genre* instead of *beige*. The choice of *beige* was influenced by uncertainties surrounding the Italian pronunciation of *genre*. To avoid any potential pronunciation issues or mistakes during the test and to align with the test's objectives, a more commonly known word among Italians that they are generally familiar with was selected.

The number of participants in this study was limited and not evenly distributed, with 18 individuals from Florence and 20 individuals from outside of Tuscany. It should be noted that the experiments were conducted online, with participants carrying out the tasks independently. However, there is uncertainty regarding whether they were indeed alone and in a quiet environment.

The selection criteria for participants also had limitations. The only strict requirement was that they belonged to either the Florentine or non-Tuscan group, and that they spoke Florentine dialect (excluding other Tuscan dialects) or any variant of Standard Italian. Additionally, participants were expected to have some level of English knowledge and usage, but there were no specific requirements regarding the extent of their English exposure or usage. As a result, the 38 participants had varying levels of English proficiency, contradicting previous literature that suggests a strong correlation between experience and the perception of L2 sounds. However, it is worth mentioning that all participants had studied English in the Italian schooling system from a young age, typically starting around 5 years old. Additionally, there were no significant differences in the amount of English used by the participants.

Future research should be conducted to examine the impact of a phonological inventory on the perception of specific sounds in words, rather than just syllables, across all possible positions within the word. It would be highly valuable to investigate this phenomenon specifically in Italian learners of English as a second language, focusing on the

Florentine dialect and the same target sounds that I have focused on. I hope my research will serve as a source of inspiration, if not a foundation, for future research.

References

Avery, P., & Ehrlich, S. (1992). *Teaching American Pronunciation*. Oxford:Oxford University Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1). https://doi. org/10.18637/jss.v067.i01

Berti, S. (2015). *Contrastive Analysis of the Italian and English Consonant Phoneme Systems.* [Master's Thesis, Università Ca' Foscari Venezia].

Bertinetto, P. M., & Loporcaro, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetic Association*, *35*(2), 131–151. http://www.jstor.org/stable/44526359

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issue in cross-language research,* (pp. 171–204). Baltimore, MD: York Press.

Boersma, P., & Weenink, D. (2023). Praat: doing phonetics by computer [Computer program]. Version 6.2.06, retrieved 6 February 2023 from https://www.praat.org

Bohn, O.-S. (2002). On phonetic similarity. In P. Burmeister, T. Piske & A. Rohde (Eds.), *An integrated view of language development. Papers in honor of Henning Wode*, (pp. 191-216). Trier: Wissenschaftlicher Verlag.

Bohn, O.-S. (2018). Cross-language and second language speech perception. In E. M. Fernandez & H. S. Cairns (Eds.), *The handbook of psycholinguistics*, (pp. 213-239). Hoboken, NJ: Wiley Blackwell.

Bohn, O.-S. (2023). *Predicting difficult difficulty. Perceptual similarity* [PowerPoint slides]. Department of English, Aarhus Universitetet.

Bohn, O.-S., & Best, C. T. (2012). Native-language phonetic and phonological influences on perception of American English approximants by Danish and German listeners. *Journal of Phonetics,* 40(1), (pp. 109-128).

Bohn, O.-S-, Best, C.T., Avesani, C., &Vayra, M. (2011). Perceiving through the Lens of Native Phonetics: Italian and Danish Listeners' Perception of English Consonant Contrasts. *International Congress of Phonetic Sciences*.

Bohn, O.-S., &Steinlen, A. K. (2003). Consonantal context affects cross-language perception of vowels. In *Proceedings of the 15th International Congress of Phonetic Sciences*, (pp. 2289-2292).

Brannen, K. (2002). The Role of Perception in Differential Substitution. *Canadian Journal of Linguistics/Revue canadienne de linguistique, 47*, 1 - 46.

Calamai, S. (2017). Tuscan between standard and vernacular: a sociophonetic perspective. In Cerruti, M., Crocco, C., & Marzo, S. (Ed.), *Towards a New Standard: Theoretical and Empirical Studies on the Restandardization of Italian*, (pp. 213-241). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9781614518839-008

Chang, C.B. (2019). The phonetics of second language learning and bilingualism. *The Routledge Handbook of Phonetics*.

De Mauro, T. (1972). *Storia linguisticadell'Italiaunita*. Bari:Laterza.

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). Baltimore, MD: York Press.

Flege, J. E. (2005). Origins and development of the Speech Learning Model. DOI: 10.13140/RG.2.2.10181.19681

Flege, J. E. & Bohn, O.-S. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3-83). Cambridge, UK: Cambridge University Press, .

Flege, J.E., MacKay, I.R., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America, 106 5*, (pp. 2973-2987).

Giannelli, L. (2000). *Toscana*. Pisa: Pacini.

Giannelli, L., & Savoia, L. M. (1978). L'indebolimento consonantico in Toscana, I. *Rivista Italiana di Dialettologia* 2, (pp. 25-58).

Guion, S.G., Flege J.E., Akahane-Yamada R., & Pruitt J.C. (2000). An investigation of current models of second language speech perception: the case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America,* 107(5), (pp. 2711-2724). DOI: 10.1121/1.428657. PMID: 10830393.

Horslund, C. S., & Bohn, O.-S. (2022). Assimilation patterns predict L2 identification accuracy of English initial consonants. *Journal of Second Language Pronunciation*, 8, (pp. 218-247).

Johnson, G., & Maydee, A. (2012). Dialectal Allophonic Variation in L2 Pronunciation.

Ladefoged, P. (1990). Some reflections on the IPA. *Journal of Phonetics*, 18, (pp. 335-346).

Lass, R. (1984). *Phonology. An introduction to basic concepts*. Cambridge: CUP.

Marotta, G. (2008). Lenition in Tuscan Italian (Gorgia Toscana). In J. Brandão de Carvalho, T. Scheer & P. Ségéral (Ed.), *Lenition and Fortition* (pp. 235-272). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110211443.2.235

McGuire, G.L. (2010). Perception A Brief Primer on Experimental Designs for Speech Perception Research.

Miller, G.A., & Nicely, P.E. (1955). An Analysis of Perceptual Confusions Among Some English Consonants. *Journal of the Acoustical Society of America, 27*, (pp. 338-352).

Paradis, C., &Lacharité, D. (2012). The Influence of Attitude on the Treatment of Interdentals in Loanwords : Ill-performed Importations. *Catalan journal of linguistics, 11*, (pp. 97-126).

Pavlovia. (n.d.). Retrieved 4th November 2023, from https://pavlovia.org/

Polka L. (1991). Cross-language speech perception in adults: phonemic, phonetic, and acoustic contributions. *The Journal of the Acoustical Society of America*, *89*(6), 2961–2977. https://doi.org/10.1121/1.400734

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy.*Behavior Research Methods.* 10.3758/s13428-018-01193-y

Pisoni, D.B. (2018). Speech Perception: Research, Theory, and Clinical Application. In Fernández, E.M., & Cairns, H.S. (Eds.), *The Handbook of Psycholinguistics*, (pp. 203-212). https://doi.org/10.1002/9781118829516.ch9

R Core Team (2018). R: A language and environment for statistical computing. *R Foundation for Statistical Computing.* Vienna, Austria. https://www.R-project.org/

Reis, M.S., Kluge, D.C., & Bettoni-Techio, M. (2008). Different Task Type and the Perception of the English Interdental Fricatives.

Schmidt, A. M. (2007). Cross-language consonant identification. In Bohn, O.-S., & Munro, M. J. (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege,* (pp. 185-200. John Benjamins Publishing Company.

Sebastián-Gallés, N., & Díaz, B. (2012). First and Second Language Speech Perception: Graded Learning. *Language Learning*, 62, (pp.131-147). https://doi.org/10.1111/j.1467-9922.2012.00709.x

Shannon, R. V., Jensvold, A., Padilla, M., Robert, M. E., & Wang, X. (1999). Consonant recordings for speech testing. *The Journal of the Acoustical Society of America*, *106*(6). https://doi.org/10.1121/1.428150

Tabain, M. (1998). Non-sibilant fricatives in English: spectral information above 10 kHz. *Phonetica*, *55*(3), (pp. 107–130). https://doi.org/10.1159/000028427

Thomas, M. (2013). History of the study of second language acquisition. In Herschensohn, J., & Young-Scholten, M. (Eds.) *The Cambridge Handbook of Second Language Acquisition*, (pp. 26-45). Cambridge University Press.

Werker, J.F. (1995). Age-Related Changes in Cross-Language Speech Perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research,* 155-169. Baltimore, MD: York Press.

Wheelock, A. (2016). Phonological Difficulties Encountered by Italian Learners of English: An Error Analysis.

Yeni-Komshian, G. H. (1998). Speech perception. In J. B. Gleason & N. B. Ratner (Eds.), *Psycholinguistics*. (pp. 107- 156).

Appendix A - Language Profile Questionnaire

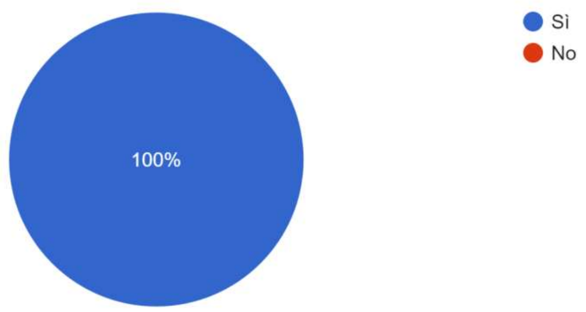

Figure A1

Graph reporting the total responses to the question "Sei di madrelinguaitaliana?"("Is Italian your mother tongue?"). The answers were: yes/no.

Figure A2

Graph reporting the total responses to the question "Quando parli con altrepersone di altre parti d'Italia, riesci a percepiredifferenzenel modo di parlare?"("Are you able to detect differences in the way other people speak, when you have a conversation with them?). The responses where:

- *Yes, I can spot the differences, and I can also identify the dialects*
- *Yes, I can spot the differences, but I can't name the different dialects*
- *No, I rarely perceive differences*
- *No, I didn't know anything about that*

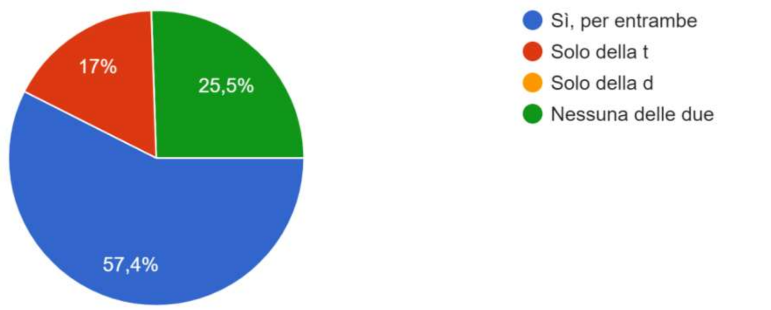

Figure A3

Graph reporting the total responses to the question "Sai che cosa è l'aspirazione della t e/o della d?". ("Do you know what "aspirazione" of t and/or d sound is"). The possible answers were:

- *Yes, both*
- *Only for the t sound*
- *Only for the d sound*
- *None of them*

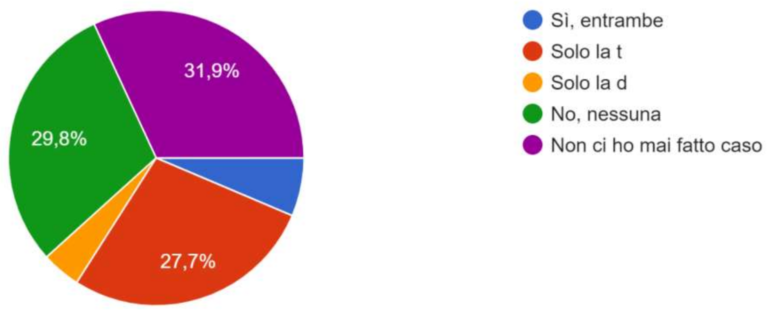

Figure A4

Graph reporting the total responses to the question "Tendi adaspirare la t e/o la d quandoparli?". ("Are you used to spirantising the t and/or d when speaking?"). Possible answers:

- *Yes, both*
- *Only the t sound*
- *Only the d sound*
- *None of them*
- *I have never noticed it*

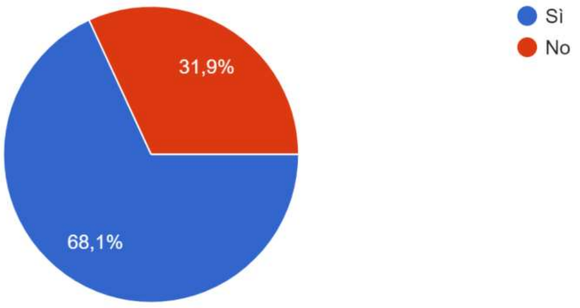

Figure A5

Graph reporting the total responses to the question "Riesci a sentire la stessa aspirazione nella parlata di altre persone?". ("Can you perceive the same spirantisation when other people are speaking?"). Answers were yes/no.

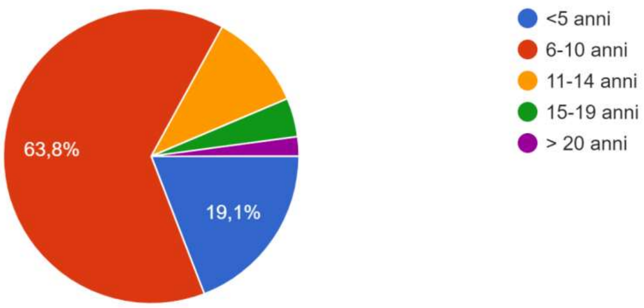

Figure A6

Graph reporting the total responses to the question "A cheetàhaiiniziatoaimpararel'inglese?". ("How old were you when you started learning English?")

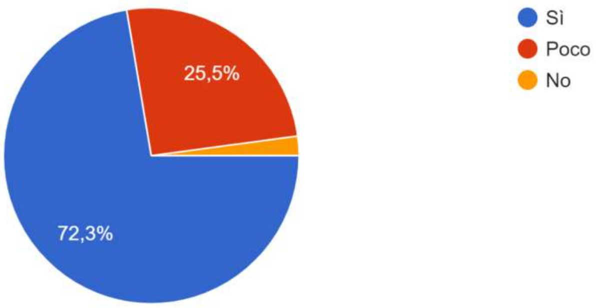

Figure A7

Graph reporting the total responses to the question "Usi tuttoral'inglese?"("Are you currently practising English?). Possible answers: yes/just a little/no.

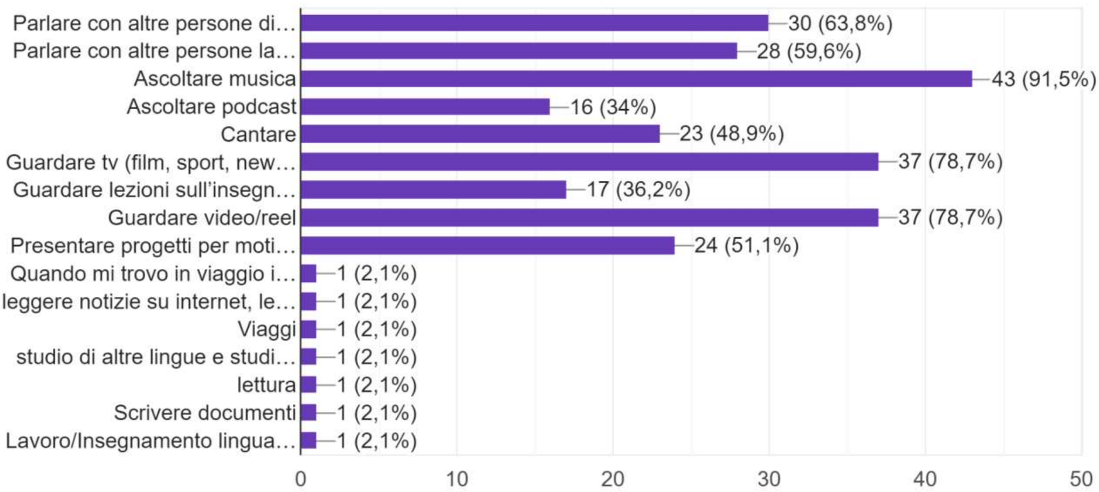

Figure A8

Graph reporting the total responses to the question "Se sì (o poco), in quali delle seguenti attività? [scegliuna o piùopzioni]. ("If you do use English (or just a little), which of the following activities fit?").

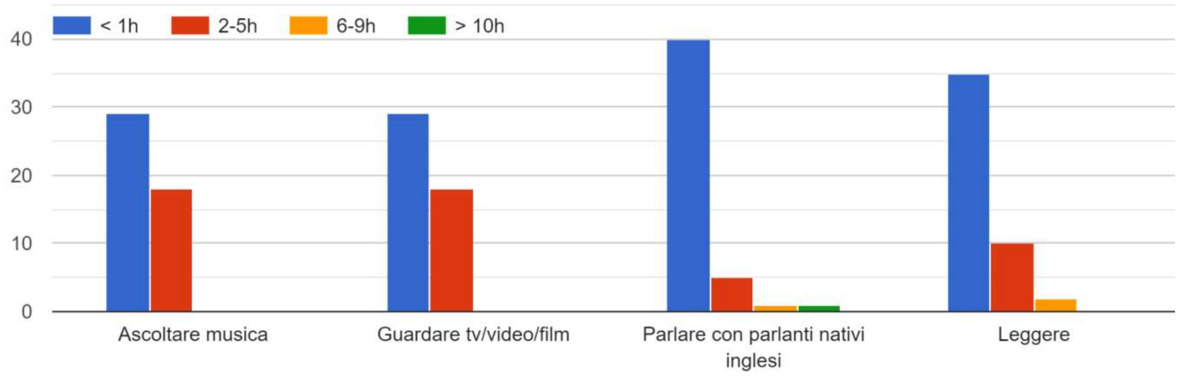

Figure A9

Graph reporting the total responses to the question "Quanto tempo spendi al giorno per le seguenti attività in lingua inglese?". ("How much time do you usually spend doing the following activities?"). Responses:

- *Listening to music*
- *Watching tv/videos/films*
- *Talking to native-English speakers*
- *Reading*

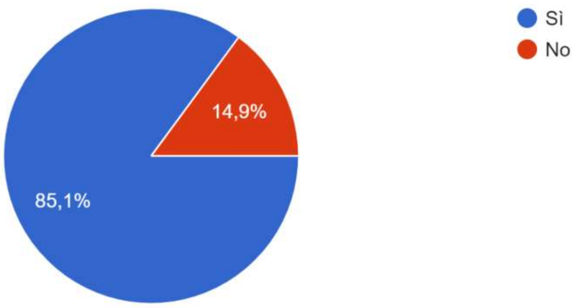

Figure A10

Graph reporting the total responses to the question "Ritieni importante avere una corretta pronuncia della lingua?". ("Do you think it is important to possess good skills in the pronunciation of a language?). Answers: yes/no.

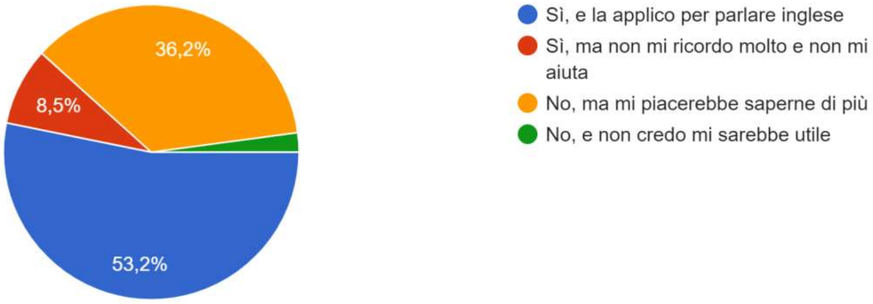

Figure A11

Graph reporting the total responses to the question "Hai mai studiato la corretta pronuncia dell'inglese?". ("Have you ever studied the correct English pronunciation?"). Answers:

- *Yes, I do apply it when speaking English*
- *Yes, but I honestly don't remember much and it is not helpful*
- *No, but I would love to know more about it*
- *No, I think it wouldn't be helpful*

Created by and taken from Google Forms [Last access 31st January 2024]

Appendix B - The experimental procedure instructions

Here I gathered together the text of the instructions for the experiment on Pavlovia.com, as was presented to the participants (note: *FASE 1* and *FASE 2* correspond to Experiment 1 and Experiment 2 respectively).

Instructions

FASE 1

Welcome!

Premi la barra spaziatrice per la fase 1, e ogni volta che vorrai procedere con le istruzioni.

Dopo aver sentito la coppia di suoni, indica quanto credi siano simili tra di loro.

In totale ascolterai 64 coppie di suoni.

Facciamo una PROVA

practice

NB: non ci sono risposte giuste o sbagliate.

Premi la barra spaziatrice quando vuoi iniziare.

64 stimuli

Well done!

Hai concluso la fase 1.

FASE 2

Premi la barra spaziatrice per la fase 2

Sentirai un suono.

Indica a quale versione ortografica lo associ, cliccando con il mouse sopra la scritta.

Le risposte ortografiche sono le seguenti:

f, v, s, z, sh, t, d, beige, THink, THat

di beiGE considera solo il suono finale, di THink e THat solo il suono iniziale

Facciamo una PROVA

practice

In totale ascolterai 48 suoni.

NB: non ci sono risposte giuste o sbagliate

Premi la barra spaziatrice quando vuoi iniziare.

48 stimuli

Hai completato anche la fase 2.

Grazie per aver partecipato!

(non chiudere la pagina)

Appendix C – Spectrograms

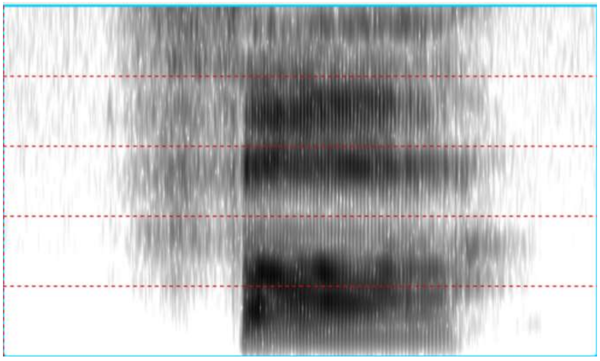

Figure C1

Spectrogram of the sound /tha/ as used in the experimental procedure. Taken from Praat(2023).

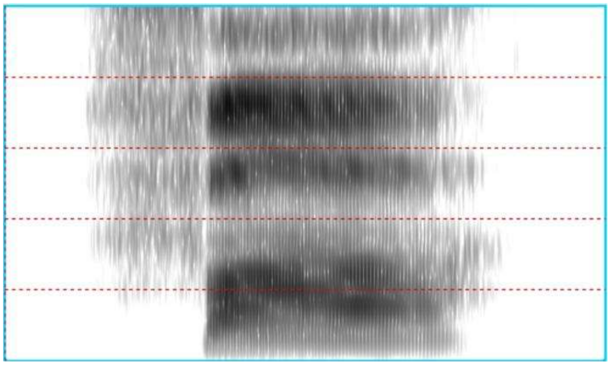

Figure C2

Spectrogram of the sound /fa/ as used in the experimental procedure. Taken from Praat (2023).

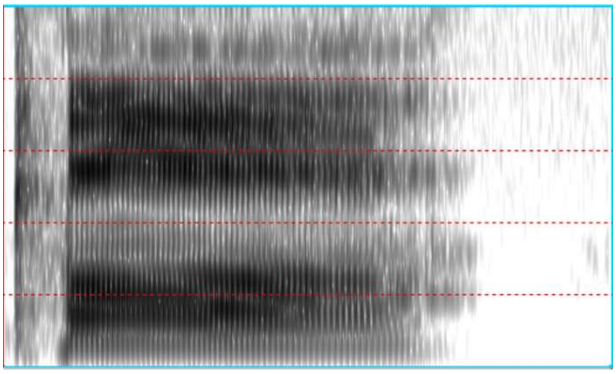

Figure C3

Spectrogram of the sound /ta/ as used in the experimental procedure. Taken from Praat (2023).

Created by and taken from Praat (2023). [Last access 11th February 2024]

Appendix D - Modulo per l'espressione del consenso informato

Perception of L2 English sounds by different groups of Italian speakers

Gentile partecipante,

Il presente studio è condotto dalla laureanda Alessia Vignoli sotto la supervisione del professor Pavel Duryagin del Dipartimento di Studi Linguistici e Culturali Comparati dell'Università Ca' Foscari di Venezia sulla piattaforma online *Pavlovia* e *Google Moduli*. Accettando questo modulo, esprime il suo consenso alla partecipazione allo studio e alle attività in esso incluse.

La partecipazione a questo studio è volontaria e potrà decidere di abbandonarlo in qualsiasi momento senza alcun tipo di conseguenza negativa. Esprimendo il suo consenso, autorizzerà i/le ricercatori/trici ad archiviare in formato digitale ed elaborare in maniera confidenziale i suoi dati personali per l'intera durata del progetto di ricerca. A tutela della sua privacy, tutti i dati raccolti non saranno mai riconducibili alla sua persona, in accordo con il codice etico e di condotta dell'Università Ca' Foscari di Venezia e con le normative vigenti. I dati verranno trattati in forma anonima in accordo con il Regolamento UE 2016/679 e il Decreto Legislativo n. 196/2003; inoltre, i risultati delle analisi dei dati verranno presentati e pubblicati in tesi, libri o articoli per riviste scientifiche in forma aggregata e anonima. Può richiedere in ogni momento di modificare, rettificare o eliminare il suo consenso alla partecipazione allo studio e tutti i dati raccolti contattando il/la responsabile della raccolta dati.

Lo studio e i moduli che le viene chiesto di compilare hanno ricevuto l'approvazione della Commissione Etica di Ateneo in data 05.02.2020, verbale n. 1/2020 (per ulteriori informazioni: commissione.etica@unive.it).

Metodologia di ricerca

Il presente studio è rivolto a soggetti di età superiore a 18 anni madrelingua italiani. L'interesse principale è quello di indagare la percezione di alcuni suoni della lingua inglese da parte di parlanti nativi italiani con accento fiorentino, in contrapposizione a parlanti italiani con accento non toscano. Lo studio è composto da due parti che verranno svolte con il proprio pc. Si raccomanda di svolgere l'esperimento in un ambiente isolato e il più possibile esente da rumori esterni.

La prima parte consiste nella compilazione di un questionario sul profilo linguistico, che potrà essere eseguito dopo aver accettato di partecipare allo studio, in coda a questo Google form. La compilazione del questionario le richiederà circa 5 minuti.

La seconda parte viene svolta sulla piattaforma online *Pavlovia*, ed è composta da due task: nella fase 1 ascolterà 64 coppie di suoni e le verrà chiesto di giudicare quanto si somigliano i due suoni proposti. Nella fase 2 ascolterà 48 suoni isolati, e per ognuno dovrà indicare il suono che le sembra di aver sentito, scegliendo tra le risposte ortografiche che troverà sullo schermo. La seconda parte ha una durata di circa 15 minuti.

Contatti

Per qualsiasi domanda relativa alle procedure dello studio e per modificare/revocare il consenso alla partecipazione allo studio, ora o in futuro, può contattare:

- Supervisore della ricerca: Professor Pavel Duryagin.

Email: pavel.duryagin@unive.it; telefono: 041 234 9491;

- Ricercatore/responsabile della raccolta dati: laureanda Alessia Vignoli.

Email: 893467@stud.unive.it;

- Eventuali altri recapiti: Staff BemboLab.

Email: bembolab@unive.it; telefono: 041 234 5738 / 041 2345748.

Informativa sul trattamento dei dati nell'ambito del progetto *Perception of L2 English sounds by different groups of Italian speakers*

ai sensi dell'art.13 del Regolamento UE 2016/679 ("Regolamento")

Con il presente documento, l'Università Ca' Foscari Venezia ("Università") le fornisce informazioni in merito al trattamento dei dati personali raccolti all'interno del progetto di tesi magistrale denominato *Perception of L2 English sounds by different groups of Italian speakers* che si prefigge di indagare la percezione dei suoni della lingua inglese da parte di diversi gruppi di parlanti nativi italiani. Il progetto è condotto dalla laureanda Alessia Vignoli e supervisionato dal professor Pavel Duryagin quale Principal Investigator. Ove necessitasse di ulteriori informazioni relative al progetto, la preghiamo di contattare il Principal Investigator scrivendo all'indirizzo di posta elettronica pavel.duryagin@unive.it.

Il progetto è stato redatto conformemente agli standard metodologici del settore disciplinare interessato ed è depositato presso il Laboratorio BemboLab – Dipartimento di Studi Linguistici e Culturali Comparati dell'Università Ca' Foscari Venezia ove verrà conservato per cinque anni dalla conclusione programmata della ricerca stessa.

1. Titolare del Trattamento

Il Titolare del Trattamento è l'Università Ca' Foscari Venezia con sede legale in Dorsoduro 3246, 30123 Venezia, rappresentata dal Magnifico Rettore *pro tempore*.

2. ResponsabiledellaProtezionedei Dati

L'Università Ca' Foscari ha nominato il "Responsabile della Protezione dei Dati", che può essere contattato scrivendo all'indirizzo di posta elettronica dpo@unive.it o al seguente indirizzo: Università Ca' Foscari Venezia, Responsabile della Protezione dei Dati, Dorsoduro 3246, 30123 Venezia (VE).

3. Categorie di Dati Personali, Finalità e Base Giuridica

Il trattamento ha ad oggetto i seguenti dati personali del partecipante: dati anagrafici (nome, cognome, età), di contatto (indirizzo email), esperienze accademiche e/o professionali.

I predetti dati saranno raccolti attraverso questo Google Form.

Il trattamento dei dati personali verrà effettuato con strumenti cartacei ed informatici, adottando misure tecniche e organizzative adeguate a proteggerli da accessi non autorizzati o illeciti, dalla distruzione, dalla perdita di integrità e riservatezza, anche accidentali.

Per la tutela della riservatezza dei partecipanti, i dati verranno successivamente privati dei riferimenti direttamente identificativi (ad es. nome e cognome, codice fiscale, etc.), in modo che non siano più immediatamente riconducibili al soggetto a cui si riferiscono, e analizzati ai soli fini della realizzazione del suddetto progetto.

Le attività di ricerca sono svolte nell'ambito dell'esecuzione delle finalità istituzionali di ricerca scientifica dell'Ateneo, pertanto la base giuridica è rappresentata dall'art. 6.1.e) del Regolamento ("esecuzione di un compito di interesse pubblico"). Lei potrà revocare il suo consenso in qualsiasi momento senza subire alcun pregiudizio, scrivendo al Responsabile della Protezione dei Dati personali ai recapiti sopra indicati. L'Ateneo si asterrà dal trattare ulteriormente i predetti dati personali salvo sussistano motivi cogenti che legittimino la prosecuzione dello stesso.

4. Tempi di Conservazione

I dati saranno conservati per la durata del progetto e fino a un massimo di 5 anni, e successivamente cancellati.

5. Destinatari e Categorie di Destinatari dei Dati Personali

I dati raccolti saranno trattati dai ricercatori dell'Università e dai ricercatori impegnati nel progetto, che agiscono sulla base di specifiche istruzioni fornite in ordine alle finalità e modalità del trattamento medesimo, nonché da soggetti che forniscono servizi ausiliari all'Università nominati 'responsabili del trattamento'. La lista aggiornata dei responsabili del trattamento è disponibile alla pagina: https://www.unive.it/pag/34666/.

I dati, in forma aggregata ed anonima (in modo da non renderla identificabile), potranno inoltre essere comunicati ad altre Università o enti per lo svolgimento delle attività di ricerca e diffusi per attività di disseminazione dei risultati (ad es. in pubblicazioni, rapporti di ricerca, banche dati nonché citazioni durante lezioni, seminari e convegni). Potranno altresì esaminare tutta la documentazione (comprensiva dei dati identificativi dei partecipanti) raccolta nell'ambito del progetto sia organismi nazionali e internazionali sia comitati delle riviste scientifiche italiane e straniere al fine di controllare che la ricerca sia condotta correttamente e in conformità alle disposizioni vigenti, nonché eventuali auditor.

6. Diritti dell'Interessato e Modalità di Esercizio

Lei potrà esercitare nei confronti dell'Università tutti i diritti previsti dagli artt. 15 e ss. del Regolamento; in particolare, potrà ottenere: l'accesso ai dati personali, la loro rettifica o integrazione, la cancellazione (c.d. "diritto all'oblio"), la limitazione e l'opposizione del trattamento. La richiesta potrà essere presentata, senza alcuna formalità, contattando direttamente il Principal Investigator professor Pavel Duryagin (pavel.duryagin@unive.it) e/o il Responsabile della Protezione dei Dati all'indirizzo dpo@unive.it ovvero inviando una comunicazione al seguente recapito: Università Ca' Foscari Venezia – Responsabile della Protezione dei dati, Dorsoduro 3246, 30123 Venezia. In alternativa, è possibile contattare l'Università, scrivendo a PEC protocollo@pec.unive.it.

Inoltre, se ritiene che i dati personali siano stati trattati in violazione a quanto disposto dal Regolamento, potrà fare reclamo al Garante per la Protezione dei Dati Personali o adire le opportune sedi giudiziarie.

Il/La sottoscritto/a (nome e cognome)

dichiara

di aver letto con attenzione e compreso le informazioni contenute nel presente documento. Dichiara di esprimere il proprio consenso a partecipare allo studio qui descritto e dichiara di *aver letto l'informativa sul trattamento dei dati personali. Il consenso potrà essere modificato/revocato in qualsiasi momento.*

Riceverà una copia del modulo di consenso informato compilato al termine della compilazione.

- Acconsento a partecipare allo studio e dichiaro di aver letto l'informativa sul trattamento dei dati
- Non acconsento a partecipare allo studio e dichiaro di aver letto l'informativa sul trattamento dei dati