Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Magistrale in Linguistica
Classe LM-39

Tesi di Laurea

Vowel Reduction Phenomena in L2 Italian

Relatore
Prof. Laura Vanelli
Prof. Leendert Plug

Laureando
Caterina Pace
n° matr.1132245 / LMLIN

Anno Accademico 2016 / 2017

Contents

List of Figures

List of Tables

Acknowledgements

As this dissertation represents the end of a journey that I didn't undertake alone, there is a long list of people that I want to thank.
In the first place, I would like to thank professor Laura Vanelli, who supervised me during the preparation of this dissertation with grace and a considerable amount of patience, and professor Leendert Plug, who helped me greatly before and after the experiment.
Another thank you goes to all the professors of both the University of Leeds and the University of Padua, who offered support in many little ways, as professors Cécile De Cat, Clare Wright and many others, and who taught me what I know today.

I am also grateful to Elisabetta, who was kind enough to tolerate me and all my questions in the months of the work in process, and all the friends in Padua and Leeds that have showed me support when I most needed it. Without you, my two years as a graduate student would have not been as joyous and peaceful as they have been.

A last heartfelt thank you goes to my parents, who, against all the odds and obstacles that they had to face in the last two years, were nothing but supportive, and my uncle Antonio, aunt Antonella and cousin Gianstefano, my family in the north, who were incredibly gracious in offering me another place to call home and never complaining.

Italian Abstract

La riduzione vocalica è per definizione la centralizzazione delle vocali in posizione non accentata (Kingston, 2007, p. 418).
Essa è generalmente distinta in fonologica e fonetica; mentre nel primo caso sarebbe un processo obbligatorio dipendente da fattori fonologici e morfologici, nel secondo caso dipenderebbe da modificazioni del contesto (Harrington, 2010, p.91).

Diversi autori offrono diverse interpretazioni della riduzione vocalica fonologica. Flemming (2004) ritiene che dipenda dal rapporto di due diverse costrizioni, una che promuove la distintività dei suoni richiedendo una certa distanza minima l'uno dall'altro e l'altra che ne favorisce i contrasti, alle quali se ne aggiunge una ulteriore che opera in relazione all'accento e provoca quindi i cambiamenti nel sistema vocalico. Crosswhite (2004) invece propone l'esistenza di una riduzione vocalica che incrementi i contrasti (come esemplifica attraverso l'esempio della diminuzione dell'inventario vocalico italiano da sette a cinque vocali in posizione non accentata) e una che invece incrementi la prominenza, semplicemente alzando la vocale bassa, da cui scaturisce l'innalzamento su tutto il sistema vocalico, ad una posizione medio-centrale.

Anche a livello fonetico la riduzione vocalica è piuttosto discussa. Mentre per Lindblom (1963, e studi successivi) essa è dipendente dal fenomeno di undershoot, per van Bergem (1993) è invece il risultato di una serie di fattori che operano in modo diverso rispetto a quanto invece sostenuto da Lindblom.
Più nel dettaglio, Lindblom (1963) sostiene che, essendo la vocale non accentata più corta rispetto alla vocale accentata, a parità di forza e velocità degli articolatori, essa non sarà pronunciata con la stessa cura della vocale accentata; al contrario, proprio a causa della diminuzione del tempo disponibile, gli articolatori non riescono a raggiungere la posizione target per la sua produzione

canonica (undershoot). Studi successivi di Lindblom hanno risposto alle critiche mosse soprattutto all'idea che forza e velocità degli articolatori rimangano sempre uguali ideando il modello H&H (hyper and hypoarticulation), che, basandosi sul concetto che l'obiettivo dei parlanti sia quello di produrre enunciati quanto più intelligibili con il minor sforzo possibile, afferma che la riduzione vocalica sia dunque una strategia di diminuizione dello sforzo che si applica in relazione allo stile del linguaggio.
D'altro canto van Bergem (1993) cita tra i vari fattori che causano la riduzione vocalica l'assenza dell'accento, che è responsabile di una produzione meno accurata della vocale a prescindere dalla sua durata, il contesto, che influisce sulla vocale a seconda della sua 'compatibilità' con essa, la categoria e la frequenza della parola, l'intonazione della frase e lo stile del linguaggio. In particolare, le vocali delle parole funzione monosillabiche sono articolate con meno precisione delle vocali nelle parole contenuto a prescindere; similarmente, le vocali nelle parole più frequenti sono articolate meno chiaramente delle vocali nelle parole più frequenti.

In inglese la riduzione vocalica senza dubbio di natura fonologica ed è associata generalmente da uno schwa (ə), sebbene esistano numerose variazioni. Questo è probabilmente dovuto al fatto che schwa è la vocale più frequente dell'inventario inglese (Roach, 2009), data la sua capacità di sostituire qualsiasi vocale inglese purché in posizione non accentata (Yavas, 2011, p. 88). Esso tende ad essere sempre presente nelle parole funzione, mentre le parole contenuto devono avere almeno due o più sillabe (Stevens, 1998, p. 574) perché ci possa essere uno schwa. Per l'italiano invece la situazione è più complicata. Solo di recente, studi soprattutto ad opera di Savy hanno provato che la riduzione vocalica avviene anche in italiano tramite dati ottuenuti da un'analisi spettroacustica del parlato formale e connesso.

In vista dell'esperimento e della discussione dei risultati, le teorie principali sull'acquisizione della fonetica di una seconda lingua sono state riviste. La teoria di Selinker (1972) dell'Interlingua si pone in contrasto con la Contrastive Analysis Hypothesis di Lado (1957) in riferimento al ruolo della L1: mentre la prima ritiene che gli apprendenti costruiscanola loro versione della L2 e tratta gli errori di pronuncia di conseguenza, la seconda li spiega in termini di transfer da L1. Altre teorie, come l'Onthogeny Model di Major (1987), la marcatezza tipolog-

ica di Greenberg (1976) e gli studi sulla percezione, ampliano e migliorano la conoscenza attuale dell'acquisizione di una seconda lingua.
In quanto all'acquisizione della riduzione vocalica, probabilmente come conseguenza del fatto che sia gli studi di acquisizione della fonetica italiana sia la scoperta della riduzione vocalica in italiano sono piuttosto recenti, pare non esistano studi riguardo all'italiano.
L'esperimento si propone dunque come modo per riempire questo vuoto. Sebbene la riduzione vocalica esista sia in italiano che in inglese, essa avviene in modo differente nelle due lingue; resta dunque da osservare cosa succeda nell'interlingua. Le possibilità sono due: o si verifica un transfer dalla lingua inglese, per cui la riduzione vocalica avrà le stesse caratteristiche della L1, oppure non si verifica transfer, e dunque la riduzione vocalica avrà caratteristiche discordanti.

L'obiettivo dell'esperimento è la descrizione e analisi della vocale /a/ in posizione accentata, non accentata in posizione finale e non accentata in posizione interna alla parola.
A questo scopo, cinque studenti inglesi, tre ragazze e due ragazzi, principianti di italiano, tra i 18 e 20 anni, sono stati scelti per la registrazione di 24 parole scelte appositamente. Sono stati dunque fatti accomodare in uno studio di registrazione e sono state mostrate loro delle immagini (corrispondenti alle parole scelte) di cui dovevano pronunciare il nome in italiano.
La scelta di compito di denominazione delle immagini, mai applicato in precedenza in questo tipo di studi, è motivata da una serie di fattori che controbilanciano il fatto che non sia naturale: è infatti perfetto per apprendenti principianti che non sono ancora a loro agio nell'uso della lingua e quindi non pronti a sostenere una conversazione, ha il merito di permettere il controllo dell'output e garantire una maggiore omogeneità dei dati e di evitare effetti negativi dovuti alla fatica.
Le parole sono state scelte in modo che contenessero tutte almeno una /a/, che fossero trisillabiche e che fossero conosciute dai partecipanti. A questo scopo, le parole sono state selezionate dal loro libro di testo. Gli stimoli così scelti contenevano 16 /a/ accentate, 17 non accentate all'interno della parola e 9 non accentate in fine di parola.
Le registrazioni così ottenute sono state tagliate con Audacity in modo tale da creare un file audio per ciascuna parola e le vocali di interesse analizzate con Praat. Problemi nell'identificazione del punto iniziale e finale delle /a/ non hanno

permesso di analizzare la durata delle vocali stesse e le trasformazioni delle F2 dovute al contesto; tuttavia, è stato possibile ottenere i valori delle F1, F2 ed F3 senza troppi problemi.

Mentre l'analisi dei dati dei singoli partecipanti ha permesso l'identificazione e, nella maggioranza dei casi, una spiegazione dei dati divergenti da quella che appare la norma, l'analisi generale di tutti i partecipanti (divisi, ovviamente, per sesso) ha portato a risultati inaspettati.
I dati, mostrati nei grafici 4.12 e 4.13, mostrano un grado di riduzione vocalica diverso sia dalla L1 dei partecipanti, l'inglese, che dalla lingua target, l'italiano. Gli insiemi che rappresentano le vocali accentate, non accentate all'interno della parola e non accentate in fine di parola, mostrano un'intersezione di gran lunga maggiore. Pur risultando dai test statistici (t-test) che la differenza tra vocali accentate e non accentate è significativa (e dunque confermando che c'è effettivamente riduzione vocalica nell'interlingua), c'è una grossa zona di sovrapposizione nelle F1, le frequenze quindi che sono maggiormente condizionate dalla riduzione vocalica.
In aggiunta, diversamente dai dati presenti in inglese e in italiano che indicano una certa differenza tra vocali non accentate all'interno della parola e non accentate in fine di parola, nell'interlingua la differenza è assolutamente non significativa e sembra addirittura scomparire.

Alcune considerazioni seguono questi risultati.
Si può ipotizzare infatti che la differenza dei risultati rispetto alle lingue interessate dipenda dal modo in cui è stato scelto di raccogliere i dati, oddia dalla denominazione di immagini. Come conseguenza alla mancanza di fonti che possano avvalorare o contraddire questa ipotesi, si propone dunque uno studio addizionale con uso di un gruppo di controllo.
Una seconda ipotesi riguarda la possibilità che i partecipanti siano stati estremamente attenti alla loro pronuncia nel corso dell'esperimento, diminuendo dunque la portata della riduzione. Tuttavia, essendo già state spiegate alcune divergenze di dati nei singoli partecipanti ipotizzando una maggiore attenzione alla pronuncia, non è logicamente possibile che le due spiegazioni possano coesistere. Dunque o le spiegazioni dei dati divergenti sono errate, oppure lo è l'ipotesi di iperattenzione nella pronuncia di tutte le parole.
Un'ultima ipotesi invece propone che i risultati riflettano una prima fase dell'acquisizione

della fonetica italiana, in cui nè la L1 nè la lingua target influenzano l'interlingua. Ne consegue che la Contrastive Analysis Hypothesis sia errata. Uno studio che includa il resto delle vocali italiane e un metodo di raccolta dati più naturale potrebbe fornire ulteriori prove.

Sebbene l'esperimento non sia esente da diverse problematiche, quali i limiti dovuti al tempo e alle difficoltà incontrate nell'analisi, esso apre una nuova serie di quesiti che, se trovassero risposta, potrebbero fornire nuove prove non solo riguardo al funzionamento della riduzione vocalica nell'interlingua, ma soprattutto riguardo al funzionamento della fonetica nell'interlingua.

Chapter 1

Introduction

Vowel reduction is a phonetic and phonological phenomenon that involves the centralization of the unstressed vowels.
Although this definition seems simple and clear-cut, under its surface there is a great number of studies and articles that, to this day, try to not only explain how it happens, but also what exactly vowel reduction is.

This dissertation is the result of a plunge into the sea of these studies to see what happens when this apparently simple phenomenon is inserted into the still evolving field of Second Language Acquisition (SLA).
As such, there is the hope for this study to become a part of the ever-growing amount of data and evidence that plays a vital role in defining what we know today and we will learn in the future years about L2 phonetic acquisition. This is especially true for the Italian studies, which are still quite new when compared to the English ones.
Thus, while this work started as the desire to satisfy a deeply-ingrained curiosity about the inner workings of phonetics and SLA, the driving force behind the dissertation is the ambition to create a new brick that may lead to a future advancement of the linguistic studies.

In order to understand vowel reduction and SLA, identify the current status of the studies and the gaps that still need to be filled, a through background review had to be done.
In chapter 2, vowel reduction is firstly explained both as a phonetic and a phonological phenomenon, while trying to be as exhaustive as possible when faced with a not irrelevant number of contrasting studies.

Then, in sections 2.2.3 and 2.2.4, vowel reduction is described as it occurs respectively in English and in Italian. This less theoretical, more practical overview is vital in the general structure of the dissertation, as it offers the basis mostly in terms of frequencies for the discussion of the experiment.
The following summary on the main theories of phonetic acquisition (section 2.3.1) similarly offers a theoretical background, that will be vital in the discussion of the experiment. Section 2.3.2 instead looks at the studies relative to vowel reduction in SLA in order to identify a gap in the literature.

Chapters 3 and 4 dive into the more practical side of the dissertation, respectively explaining and justifying the chosen methodology and showing and discussing the results.
The analysis of vowel reduction in L2 Italian of L1 English speakers showed some unexpected patterns both for the individual participants and especially in general. These patterns are discussed and in conclusion the all experiment was revised in the light of the results.

Chapter 2

Background information

2.1 Introduction

Vowel reduction is an interesting phenomenon that has been studied from many different points of view. In order to proceed to the experiment it is therefore necessary to gain an understanding on how it works and what is the state of literature focused on this topic. Therefore, in section 2.2.1 a general definition of vr and some phonological accounts of it will be given, followed, in section 2.2.2, by the main theories around the factors that originate vr and how it occurs in different languages such as English (section 2.2.3) and Italian (2.2.4).
As a look into the literature around vr acquisition would be incomplete and of difficult understanding without the proper background, an overview of the main theories of SLA will be given (2.3.1), before a brief reflection on its currents state and gaps (2.3.2).

2.2 Vowel reduction

2.2.1 What is vowel reduction

Vowel reduction can be described as the phenomenon that involves the change of vowel quality in an unstressed position (Kingston, 2007, p. 418). In other words, during the articulation of the vowel in an unstressed syllable, the tongue does not reach its target on the periphery of the vowel space, but a more *central* position. This can be more easily seen in figure 2.1, where this central area is highlighted in blue, and the common positions of the vowels are indicated at the margins of the vowel space.

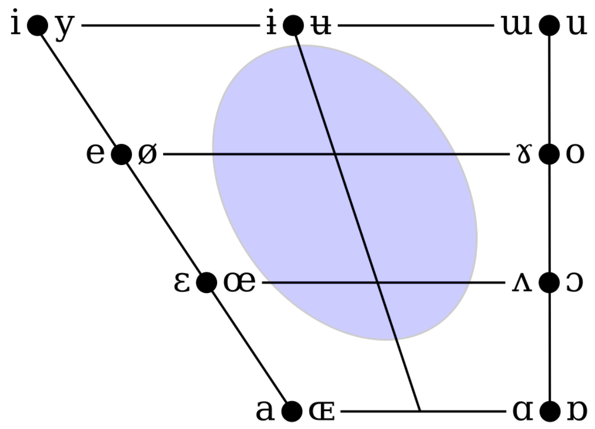


Figure 2.1: Vowel chart showing the central positions of the vowel space.

Source: `https://en.wikipedia.org/wiki/Vowel_reduction`, based on Collins and Mees (2003)

Here is an example from Yavas (2011) from English. In the word homogeneous [homoʤ*i*niəs], the vowel /i/ occurs in a stressed syllable; however, when the accent moves, as we see in homogenize [həmɑʤənaɪz], the now unstressed /i/ is reduced to a schwa [ə][1]. This change of the vowel quality is generally known as *weakening*. Keeping the example in mind, it is however important to remember that centralization of the unstressed vowel does not always lead to a ə, but many different degrees of such.

Related to Yavas's (2011) example is the distinction between phonological and phonetic reduction, which is made by numerous authors, such as, for example, Harrington (2010), Bybee (2011) and Barnes (2006) (both cited in Rallo Fabra, 2015, p.163). As Harrington (2010, p.91) says,

> "[phonological reduction] is an obligatory process in which vowels become weak due to phonological and morphological factors, as shown by the alternation between /eɪ/ and /ə/ in "Canadian" and "Canada"

[1]Vowel reduction is not a mandatory rule in English. More about it will be said in section 2.2.3.

in most varieties of English. In [phonetic reduction], vowels are phonetically modified because of the effects of segmental and prosodic context."

Another example of phonological vowel reduction, given in Flemming (2004), is based on the Central and Southern Italian vowel inventory showed in figure 2.2.

(i)	i		u	(ii)	i		u	(iii)	i		u
	e		o		e		o			a	
	ε		ɔ			a					
		a									

Figure 2.2: The Italian vowel inventory

Source: Flemming (2004, p.244), based on Maiden (1995)

In a primary stressed syllable, the Central Italian vowel inventory is the one showed in figure 2.2i. However, in an unstressed syllable, the Central Italian distinction between mid-high and mid-low vowels is neutralised (2.2ii). On the other hand, the Southern Italian has a five vowel inventory in a stressed position (2.2ii) and is reduced to a three-vowel inventory in an unstressed position(2.2iii). So, in Central Italian the words change like this:

[e]	péska	peskáre	[ε]	bέlːo	belːíno
[o]	nóme	nomináre	[ɔ]	mɔ́lːe	molːeménte

while in Southern Italian they change like this (from Mazzola, 1976, cited in Flemming, 2004, p.244):

[e]	véni	vinímu	[o]	móri	murímu

Barnes (2002), as cited in Kingston (2007, p.420), further adds to the phonetical/phonological reduction distinction that *if* phonologization of the reduced vowel occurs, the reduction is no longer motivated by the phonetical factors that caused it (section 2.2.2).

While agreeing with Barnes (2002) on the phonetic motivation of vowel reduction, which is based on Lindblom's (1963) study (section 2.2.2), Flemming (2004)

gives a detailed phonological account in which the phonetics of vowel reduction operate. He argues that it is the result of the interaction between two constraints, one that favours the distinctiveness of the sounds by requiring a certain minimal distance between them (MINDIST) and one that instead favours the maximising of the contrasts (MAXIMISE CONTRASTS).
In the Central Italian unstressed vowel inventory, a third constraint, one that prohibits short low vowels, is added. It has the consequence of neutralising the vowels [ɛ] and [ɔ].

Crosswhite (2004) proposes instead a different contrast rather than the phonetical/phonological one. She distinguishes contrast-enhancing reduction with prominence-reducing reduction. Constraint-enhancing reduction has one example in the Italian inventory, as "the neutralization of height contrasts involving mid vowels leaves the remaining contrasting vowels farther apart" (Kingston, 2007, p.419). It mostly favours noncorner vowels, especially mid-vowels. Prominence-reducing is obtained by raising the low vowel to a mid-central position with a schwa-like quality, like in Bulgarian unstressed vowels (figure 2.3).

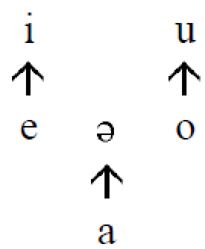


Figure 2.3: Bulgarian vowel reduction

Source: Crosswhite (2004, p.210)

The vowel /a/, Crosswhite argues, is especially favoured in this kind of reduction. The raising of the vowel has not lowers its prominence (or sonority), but also the side-effect of lowering its duration (Lehiste, 1970, cited in Crosswhite, 2004, p. 201), which is compatible with the phonetical data of vowel reduction (section 2.2.2).

In conclusion, while the general definition of vowel reduction is generally agreed upon, many theories have been made around the phonology of it. Some examples from the recent literature has been given. While Crosswhite (2004) gives a completely functional description of this phenomenon, Barnes (2006) and Flemming (2004) build their phonological account on the phonetics of it. In the next section, it will be described in more detail.

2.2.2 Phonetical explanation of vowel reduction

Vowel reduction is a complex phenomenon that occurs in many different languages. There are two main different phonetical explanations in the literature, which will now be examined.

The first study that tried to give an articulatory and acoustic explanation of it is Lindblom's (1963) "Spectrographic Study of Vowel Reduction", later followed by Moon and Lindblom (1994). Although since then many things have been disputed, it still remains an influential article regarding vowel reduction.

Lindblom's explanation revolves mainly around the concept of undershoot and duration. "Undershoot refers to a situation in which a given speech sound is articulated in a manner that does not fully instantiate the canonical realization of that sound" (Crosswhite, 2004, p. 216); in other words, it occurs when the tongue does not reach its acoustic target, mainly because it does not have enough time.

Stress has, in this reasoning, an important role. According to Ladefoged and Johnson (2011, p. 310), who mostly operate in the English phonetics field, stress is "the use of extra respiratory energy during a syllable", while, according to Kager (2007, p. 195), it is related to the concept of "prominence", where prominence is not intended as an absolute, but as "a matter of relative strength between 'stronger' and 'weaker' syllables.". In other words, there is no intrinsic value that makes a syllable prominent (or stressed) if not when in relation with other syllables. Hence, it is clear that stress is not acoustically defined by one exclusive phonetic property. There is instead a general agreement that stress is the result of multiple cues such as "higher pitch levels, longer duration, and greater loudness than unstressed syllables" (Kager, 2007, p.195).

As stress implies a longer duration of the vowel, it follows that an unstressed vowel is shorter, and this has a certain effect on the *quality* of the vowel itself. It is Lindblom's (1963) argument that vowel reduction is the effect of the shortening: as the speed of the articulatory movement is limited and thus the extent to which articulators reach their target positions depends on the timing (Lindblom, 1963),

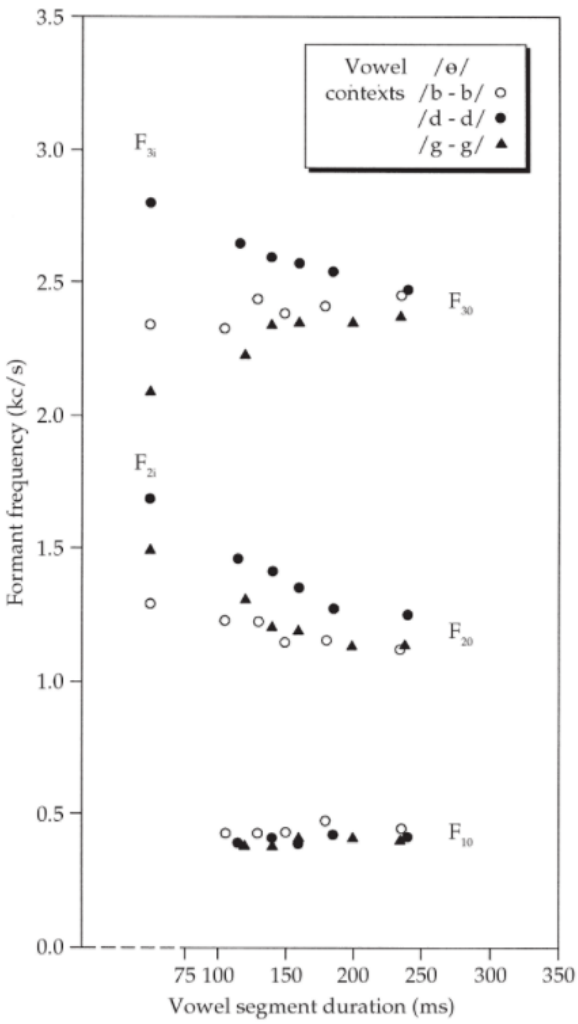


Figure 2.4: Mean values of F1, F2 and F3 plotted against vowel duration.

Source: Farnetani and Recasens (2010, p.331), based on Lindblom (1963)

it follows that a shorter vowel will be less likely to be canonically realized than a longer vowel. Being that stress has an influence on duration and that vowel reduction occurs in unstressed syllables, it is relatively easy to see a cause-effect association between stress, undershoot and vowel reduction (see figure 2.4).
The hypothesis of the significance of duration in vowel reduction is supported in the same study by a mathematical analysis of the vowel shift in different consonantal contexts, based on the idea that the articulatory effort delivered to the articulators remains constant through the utterance. This idea has been discussed by numerous authors as well as Lindblom himself. In his later studies

(Lindblom, 1990 and Lindblom, 1996) he in fact proposes the model of hyper and hypoarticulation (H&H). Segmental reduction finds an explanation under the assumption that "the aim of talkers is to produce utterances that are intelligible to listeners, but to do so with as little effort as necessary" (Hay et al., 2006, p.3022). Hence, a speaker will apply different strategies to achieve sufficient distinctiveness in their utterances, that vary in a continuum from reduced forms (hypospeech) to clear forms (hyperspeech). [2] The H&H model therefore explains vowel reductions in terms of a strategy to lower the cost of the articulation while maintaining a certain level of intelligibility and addresses the lack of a speech style variable in his previously mentioned study.

One argument has to be raised in regard to the application of the H&H model specifically in vowel reduction, as it faces the risk of becoming a vicious circle, particularly when keeping in mind the previously explained phonological explanations of vowel reduction. Altman and Carted (1989) argue that, since the possible number of vowels permitted in a stressed position is higher than the number of vowels permitted in an unstressed position, stressed syllables carry more information. For the H&H model this means that vowels in unstressed positions are more hypoarticulated (i.e. reduced) than stressed vowels. However, as vowel reduction itself leads to the neutralization of vowels in an unstressed position (therefore leading to the disparity of possible vowels in a stressed versus unstressed position), as we have seen in section 2.2.1, the circular nature of this argument appears quite evident.
A way to solve this situation without discarding completely both Lindblom's (1990) H&H model and Altman and Carted's (1989) study may be by taking into account the other factors that lead to vowel reduction.

The other main study is van Bergem's (1993) "Acoustic vowel reduction as a function of sentence accent, word stress, and word class". As other studies before his, he contradicts Lindblom by arguing that

> "there is no clear relation between the vowel separation of a speaker and the duration of his vowels. Some speakers produce vowels with relatively long durations and have a poor vowel separation, whereas

[2] For a more detailed discussion on the H&H model and how different languages enhance distinctiveness, see Hay et al. (2006).

> others produce vowels with relatively short durations and have a clear vowel separation. This is most likely dependent on the particular speech style a certain speaker uses and not on physiological constraints of his articulatory organs. In our view spectral reduction is not caused by a decreasing vowel duration, but both duration and spectral quality usually decrease at the same time if the pronunciation gets sloppier." (van Bergem, 1993, p. 21)

Instead, he claims that acoustical vowel reduction is dependent on the combination of a number of factors, which are stress, surrounding phonemes, word class, frequency of occurrence, sentence accent and the already mentioned speech style. While stress was already presented as a factor for vowel reduction in Lindblom (1963), the studies disagree on *how* it operates. While Lindblom acknowledges that an unstressed vowel is articulated with less precision, he downplays it in relation to the prominence of duration as a cause of the phenomenon; on the other hand, van Bergem emphasises the difference in the articulation precision while dismissing the duration. In addition to that, he noticed that stress can be overruled by *word class* and even more so by *frequency of occurrence*. More specifically, stressed vowels in monosyllabic function words share the same precision with unstressed vowels from content words, and vowels in frequent words are articulated less clearly than less frequent words; it follows that, being that function words are highly frequent, the word class effect may be a consequence of the frequency of occurrence effect.
While the context is not dismissed and is noticeably included in Lindblom's experiment (see figure 2.4), it plays a minor role in his theory; instead, for van Bergem it is vital when considered in conjunction with all the other mentioned factors. Vowel reduction depends on context in terms of compatibility:

> "vowels that are surrounded by 'compatible' phonemes will retain a rather high quality under various different linguistic and extralinguistic circumstances. Vowels that are surrounded by 'incompatible' phonemes, on the other hand, will show effects of contextual assimilation that can rapidly increase when the speech gets sloppier." (van Bergem, 1993, p.20)

Despite the disagreement on the importance of coarticulation in vowel reduction,

van Bergem and Lindblom agree instead of the notion that coarticulation does not always lead to centralization.
Context assimilation, as it has been said, does not operate in isolation and instead is dependent on the other mentioned factors.
Sentence accent, which is described as prominently "a sudden rise or fall in the pitch contour" (Hart et al., 1990, cited in van Bergem, 1993, p.2), belongs to a different category: if stress, context, word class and frequency of occurrence are fixed properties of words, sentence accent is only assigned temporarily to words. Although it has an effect on the reduction, it is however, compared to the other factors, of a lesser extent.

In this vowel reduction theories review, it is evident that there is a general disagreement as to how the different factors combine.
The status of vowel reduction represents another matter of importance. As the literature around vowel reduction seems mostly revolving around English and other stress-timed languages such as Swedish and Dutch, it does reflect a bias in the definition of both phonetical and phonological vowel reduction. In other words, as rhythmic differences in different languages partly rely on stress and timing (Ladefoged and Johnson, 2011), if we stand by Lindblom's (1963) theory we should expect a change as to vowel reduction works. Thus, in the next sections (2.2.3, 2.2.4) vowel reduction will be analysed as it works specifically in English and in Italian.

2.2.3 Vowel reduction in English

As most of the previously cited studies work on describing vowel reduction based on the English language, there is only a little more to add on the matter.

Acoustically speaking [ə] has some interesting qualities. While acknowledging intraspeakers variation, the F1 of [ə] may vary from a mean of 428Hz in middle position to a mean of 665Hz in word-final position for adult female speakers, while the F2 shifts accordingly to the context (Flemming and Johnson, 2007), as it is visible in figure 2.5.

The length of [ə] may vary to a minimum of less than 50ms when the consonants preceding and following it are produced with different articulators; in faster

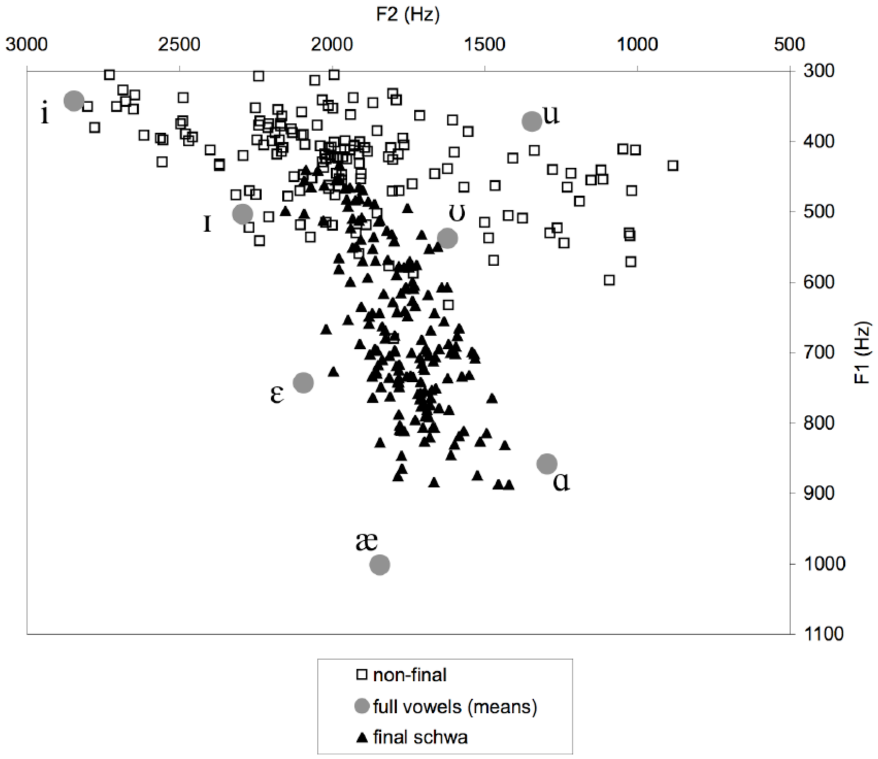


Figure 2.5: Formant frequency of final and non-final schwa from nine female English speakers.

Source: Flemming (2009)

speech, the overlap between the release of the first consonant and the closure of the following can even remove any evidence of vocalic opening (Stevens, 1998, p.576).

While the result of the reduction may vary in relation to the context, in the general literature it is mostly associated with [ə]. Although present, variations are considered not important, especially when confronted with the frequency of occurrence of [ə], which is the highest of all the English vowels (Roach, 2009, p. 65). While its occurrence is limited to unstressed syllables, the high frequency of [ə] is owed to the fact that it can substitute all the English vowels when unstressed (see the following example from Yavas, 2011, p. 88).

[i]	homogeneous [homoʤiniəs]	homogenize [həmaʤənaɪz]
[ɪ]	implicit [ɪmplɪsət]	implication [ɪmpləkeʃən]
[e]	rotate [ɹ̣otet]	rotary [ɹ̣otəɹ̣i]
[ɛ]	perpetuate [pɚpɛʧuet]	perpetuity [pɚpəʧuəti]
[æ]	enigmatic [ənɪgmætɪk]	enigma [ənɪgmə]
[ɑ]	stigmata [stɪgmɑtə]	stigma [stɪgmə]
[o]	photograph [fotəgɹ̣æf]	photography [fətɑgɹ̣əfi]
[ʌ]	confront [kənfɹ̣ʌnt]	confrontation [kɑnfɹ̣ənteʃən]
[aɪ]	design [dəzɑɪn]	designation [dɛzɪgneʃən]

Yavas (2011, p.89) also illustrates several possible variations, such as the use of the vowel [ɪ] rather than [ə] in several people's speech before palato-alveolars (for example selfish [sɛlfɪʃ]) and velars (for example metric [mɛtɹ̣ɪk]) especially in conditions of tautosyllabicity. An additional point is that, while [ə] cannot appear in stressed syllables, it is possible for an unreduced vowel to appear in a stressed syllable, as it is the case for October [ɑktəʊbe(r)], where the first syllable is not stressed.

Naturally, words in connected speech, as opposed to words in isolation, tend to accommodate more [ə] and it is especially true for function words, which, with the exception of the cases when they bear the focus of the phrase (for example, in "We can serve strawberries or grapes for dessert" "I think we should serve strawberries **and** grapes"), are always reduced and sometimes have more than one weak form (Yavas, 2011, 90). To have a weak form, content words must instead contain two or more syllables (Stevens, 1998, p. 574).

One last point to make is about the phonological nature of vowel reduction in English. It is evident in the previous example from Yavas (2011) in the alternations between full and the reduced vowel in the same morpheme. According to Fourakis (1991) and Ladefoged and Johnson (2011), phonological vowel reduction has exceptions in American English that include the monophthongs and diphthongs [ɔ, u, ʊ, ɔɪ, ɑʊ]; no similar sources could be found in relation to British English.

2.2.4 Vowel reduction in Italian

Vowel reduction in Italian is not as well-defined as in English. Some authors such as Bertinetto (1981) argue that vowel reduction does not occur; Savy and Cutugno (1997), Savy et al. (2005), Krämer (2009, as cited in Giavazzi, 2010), and the previously mentioned Flemming (2004) and Crosswhite (2004) give, on the other hand, different descriptions of the phenomenon in Italian. Savy, especially in Savy and Cutugno (1997), gives the most exhausting description, that acknowledges not only hypo and hyperspeech conditions but also diaphasic differences, offering a full picture of the situation. Unlike Flemming (2004) and Crosswhite (2004), in this description the neutralization of the mid vowels in Central and Southern Italian is not accounted for and perhaps it is treated by the authors as a regular neutralization and not as a consequence of vowel reduction.

In addiction to the presumably non-existence of neutralization as a consequence of reduction, another difference of the Italian vowel reduction lies in the [ə]. Indeed, as van Bergem (1995) states, there is a difference between languages with a phonological and non-phonological [ə]: as (generally speaking) Italian does not have a phonological [ə], it represents an important factor of discrimination. In this case, while English has generally fixed values for it (see section 2.2.3), Italian does not; to overcome the lack of those values, Savy and Cutugno (1997) makes use of the concept of centroid. By definition, a centroid is "the grand mean of all measured formant frequency of the vowel system per speaker" (Koopmans-van Beinum, 1983, p.168, cited in Savy and Cutugno, 1997, p. 182); it has the merit of not only making it possible to measure the degree of centralization, but also of representing the center of any system given its variable nature (Savy and Cutugno, 1997). The centroid is therefore a useful tool to examine different vowel systems and enabling the comparison of the data of this different systems, as Savy and Cutugno do with the regional varieties of Italian.

Although previous studies based on auditory analysis seem to find a tendency to centralization only in the meridional varieties of Italian (those varieties that *do* contain a phonological [ə]), Savy and Cutugno's (1997) spectro-acoustical analysis found that vowel reduction occurs in all the varieties. As values closer to centroid mean that a vowel is more centralized, it is then evident in figure 2.6 that there is a significant difference between stressed and unstressed vowels in all the varieties, and that unstressed vowels are indeed more centralized.

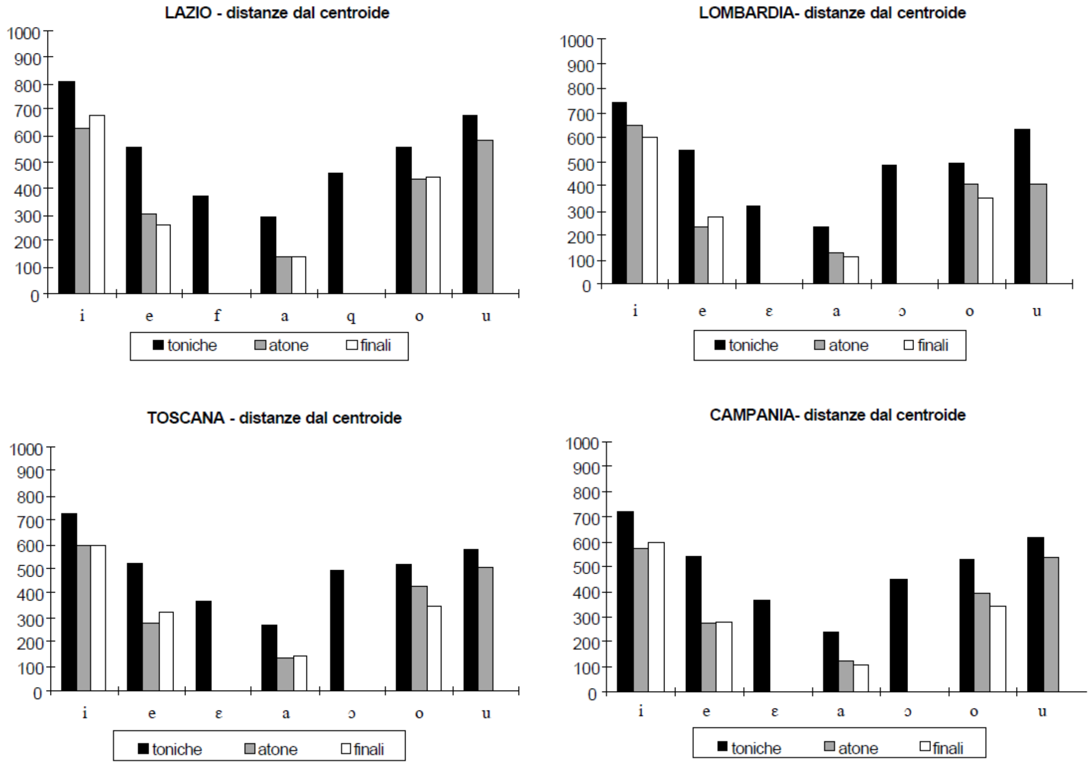
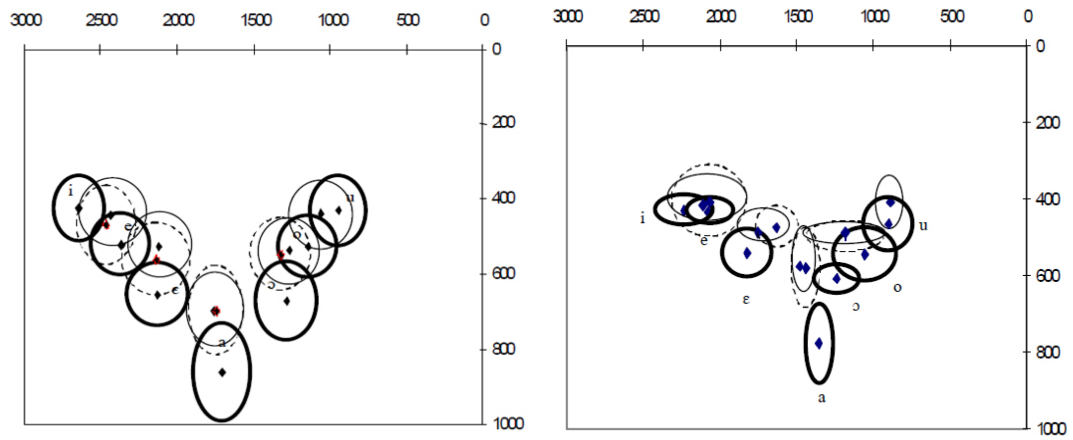


Figure 2.6: Distance from the centroid of the stressed (toniche), non-final unstressed (atone) and final unstressed (finali) vowels in the Italian varieties spoken in Lazio, Lombardy, Tuscany and Campania.

Source: Savy and Cutugno (1997, p.186)

Concentrating on the Campanian variety, Savy and Cutugno also analysed the differences between vowel reduction in what is called "parlato connesso" (PC) and "parlato spontaneo" (PSp). The acoustic segments of the PC are based on the Database on Italian Vowel Acoustics (DIVA), which contains utterances from regional news programs, characterised by high accuracy and formality (Savy and Cutugno, 1997), and can therefore be defined as hyperspeech conditions. The PSp is instead based on the Lessico di frequenza dell'Italiano Parlato (LIP, De Mauro et al., 1993), which contains spontaneous conversations, characterised by low accuracy (Savy and Cutugno, 1997), and can therefore be defined as hypospeech conditions. While vowel reduction is clearly more accentuated in hypospeech (see figure 2.7b), a certain tendency to centralization is present in hyperspeech too (see figure 2.7a), although less distinctively. In hypospeech, some tendencies are



(a) Female vocalism in hyperspeech (b) Female vocalism in hypospeech

Figure 2.7: Frequencies of Italian stressed vowels (thick line), non-final unstressed vowels (thin line) and final unstressed vowels (dotted line) in hyper and hypospeech conditions.

Source: Savy and Cutugno (1997)

clearly visible: the F2 of /i/, /e/ and /ɛ/, and /a/ tends to be lower, reduced values for F1 and overlap between the categories in the center (Savy and Cutugno, 1997). The values to which the unstressed vowels tend, Savy and Cutugno note, are those indicated for [ə].
In Giavazzi (2010, p. 344), another graph offers similar data in the Bark scale[3] (figure 2.8).

Although [ə] does not have the same morpho-phonological value as in English, the reduction of vowels does have some importance in Italian morphology. The overlap of the vowels visible in figure 2.7b indicates a decline and eventually a loss of distinction between vowels, especially in final positions. For Italian, as Savy (1999) argues, this has an impact on the inflectional suffixes and thus on the

[3]The Bark scale transforms the Hertz values so that it approximates the "frequency response of the sound-detecting hair cells of the inner ear", which are more sensitive to change in the lower frequencies than the higher frequencies. Thus, the higher frequencies are "compressed" relatively to the F1, and better reflects how vowel quality is perceived. The algorithm to convert Hz values to Bark units is:

$$z = [26.81/(1 + 1960/f)] - 0.53$$

where f is the frequency in Hertz (Watt et al., 2010, pp. 111-112).

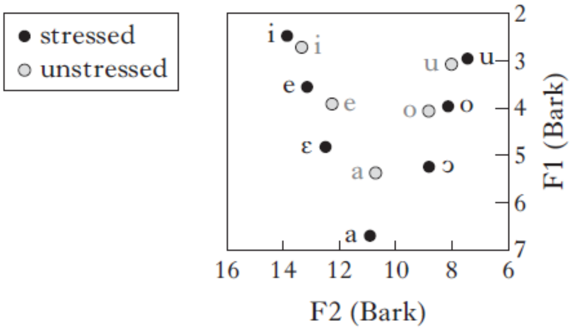


Figure 2.8: Vowel reduction in Italian, Bark scale

Source: Giavazzi (2010, p. 344), data from Albano Leoni et al. (1995)

morphological agreement. Savy found that in the noun phrase about 45% of the suffixes are not realised in spontaneous conversations, with the reduction mostly occurring in nouns and adjectives and less in articles. While several hypothesis have been raised about why the article, which is generally unstressed, keeps the vowel quantity more frequently than content words, there is no definitive explanation for this phenomenon.

In conclusion, while with characteristics different from English, several studies by Savy show that vowel reduction does occur in Italian, regardless of the spoken variety, with differences between hyper and hypospeech, with the latter having, of course, a more pronounced centralization.

2.3 Stress and vowel reduction in SLA

2.3.1 L2 phonetic and phonological acquisition: an overview

While this is not the place for a complete prospectus of the story of the studies on the acquisition of the phonetics and phonology an L2, an overview in some of the main theories in this field is however necessary for a general understanding.

One of the most widely acknowledged theories is Selinker's (1972) "Interlanguage", or Corder's (1971) "idiosyncratic dialect" or Nemser's (1971) "approximative system", which all refer to the same notion that "L2 learners construct their own version of the L2" (Eckman, 2004, p. 524). The basis of interlanguage lies in the observable patterns of a learner's speech that are not explainable via

transfer from an L1 or via input from the L2, thus responding to a major criticism moved to Lado's (1957) until then almost constantly employed Contrastive Analysis Hypothesis (Major, 1998, p. 132).
While the CAH explained pronunciation errors in terms of a L1 transfer, comparing the phonemes and their distribution within the L1 and L2, the interlanguage theory explains them by analysing the IL on the assumption that learners construct their own grammar. Interlanguages are thus autonomous systems with their own internal consistency.
A different point of view is Tarone's (1988) (cited in Mori, 2007, p. 25); the interlanguage is, in this framework, not an independent system with its own grammar, but a variety of the L2 with high dynamism.

Major's (1987) Onthogeny Model (cited in Mori, 2007, p. 26 and Eckman, 2004, p. 22), that later evolved in the Ontogeny Phylogeny Model (Major, 2001), works in the IL theory and states that the nature of sound substitutions (as a consequence of L1 transfer or as a consequence of L2 development) changes over time and in more or less formal environments; more specifically, L2 sound substitutions due to L1 transfer decrease as the learner progresses and as the speaking situation becomes more formal. Vice versa, sound substitutions due to L2 improvement increase over time and decreases in less formal environments (Eckman, 2004, p. 534). The OPM adds to the OM the concept of language universals that constrain the L2 phonology. While the statements about the nature of sound substitutions and style of speech have not been supported, the OM generally seems to hold true.

Markedness is one of the major issues in the study of language acquisition. Its definition may vary; according to the Prague School of Linguistics, in a binary opposition between linguistic representations the marked element is the one with a wider distribution, while Greenberg's (1976) theory of language universals defines *typological markedness* as follows.

> A structure X is typologically marked relative to another structure, Y, (and Y is typologically unmarked relative to X) if every language that has X also has Y, but every language that has Y does not necessarily have X. (Gundel et al., 1986, p. 108, cited in Eckman, 2004, p. 529)

The role of typological markedness in L2 phonology has been discussed and described by two main theories: the Markedness Differential Hypothesis (MDH;

Eckman, 1977) and the Structural Conformity Hypothesis (SCH; Eckman, 1991). While the MDH operates in the CAH implementing the possibility of giving a measure of difficulty on the basis of markedness in the L1-L2 differences, the SCH operates *outside* the CAH (Eckman, 2004, p. 532); the prediction of the difficulty in learning an L2 segment, which has been confirmed by numerous studies, is thus possible in different models of L2 acquisition.

Another issue is perception. Its first complexity lies in the existence of two different approaches, phonetical and phonological. In the first case, the listener merely recognises the various linguistic sounds in the signal (invariance); in the second, only the relevant characteristics are preserved (Mori, 2007, p. 28).
A second complexity regards the relationship between perception and production, it being, as Major (1998, p. 134) argues,

> [. . .] not straightforward because it has been shown that perception and production are not mirror images of one another, although the two are closely related. The significant components and factors in perception are necessarily crucial in the development of the UR [underlying representation] because an UR usually does not develop without perception. The UR then in turn provides input for production.

While it seems reasonable to expect that, similarly to the acquisition of the L1, an accurate perception of the sounds should precede their production, many studies found that the relation is not crystalline. It indeed seems that, if the learner has access to the written form of some contrast, its production may precede its perception (Eckman, 2004, p.519). Moreover, perception itself is not without difficulties: especially in the case of vowels, Mori (2007) argues, learners tend to find hard to categorise those not included in their own L1 or with different formant values. Flege and MacKay's (2004) study prove this by showing Italian's difficulty at discriminating the English pairs /ɑ/-/ʌ/, /ɛ/-/æ/and /i/-/ɪ/ and its relation to the frequency of use of the L1: more specifically low-L1-use participants could better identify the pairs instead of assimilating them to a single Italian vowel than high-L1-use participants.
Three main different models explain role of perception in L2 phonology: the Perceptual Assimilation Model (PAM) proposed by Best (1993, 1994, 1995), Hancin-Bhatt's (1994) Feature Competition Model (FCM) and Flege's (1995) Speech

Learning Model (SLM).
In short, according to the PAM, the non-native sounds are assimilated to the L1 categories in relation to the similarity of articulatory movement to the L1 sounds; phonetical properties different from the native phonological categories are not assimilated because they are not perceived (Mori, 2007, p. 29).
The FCM instead claims that more frequent features in the L1 will impact more on the perception of L2 sounds than less frequent features; one of its advantages is that Hancin-Bhatt provided an algorithm for computing feature prominence in a phonological system that enables testing (Eckman, 2004, p. 520).
The SLM, which is, according to Eckman (2004), "the model of speech perception and production that has been most influential for L2 pronunciation" (p. 520), gives an account of perception that is a consequence of the geometry of the L1 and L2; L2 learners can establish new phonetic categories for L2 sounds that are different from sounds in the L1 and can do so through their entire life.

The Universal Grammar is the last of the theories that will be explained. The application of the UG's principles to SLA is a complex matter: as Cook and Newson (1996) state, the main discussion revolve around whether UG can or cannot be accessed when learning an L2. They argue that there are three alternatives: the UG can be accessed, the UG can be accessed only through the L1 or the UG cannot be accessed. As all three models have been supported by studies, to this day it is not clear how it works in the L2. As for L2 phonology, as the base for its analysis UG considers not the phonemes but the distinctive features. Thus, if UG can be at least partially accessed, transfer phenomena can be explained as the application of the principle and parameters of the L1 in the L2 (Mori, 2007, p. 26).

In conclusion, many theories respond to the different needs of L2 acquisition. While the CAH, that has changed and been implemented with the typological markedness in the MDH, still remains influential in the analysis of L2 phonology, IL offers another plausible and widely recognised framework, that too has been implemented with the OM and later OPM. As the research is still ongoing for all the described models, while it appears that presently the IL is the most favoured theory, there is no ground to ignore them entirely.

2.3.2 Acquisition of stress and vowel reduction

Although the research around L2 phonology has generally focused more on segments and syllables than prosody, there is a large number of studies on stress (that, as we have seen, is tightly linked to vowel reduction) that seem to confirm that the L1 has a significant role in determining the IL stress; moreover, in some cases the learners seem to have constructed a stress system not deriving from both the L1 and L2 (Eckman, 2004, p. 537).

As for the specific languages, while the studies about L2 English vowel reduction are abundant, none of them could be found about L2 Italian vowel reduction. This is probably the consequence of two main reasons: L2 Italian phonetics and phonology is a relatively new field (Giacalone Ramat et al., 2014, p. 6), and, as it was stated in section 2.2.4, some authors argue that vowel reduction does not exist in Italian.
Because of this, while taking them with a grain of salt, some answers had to be looked in other languages' literature. Spanish, which is a syllable-timed language and phonetically similar to Italian, is akin to Italian in relation to vowel reduction, as only just recently the occurrence of vowel reduction has been recognised (Nadeu, 2014; Cobb and Simonet, 2015). Thus, while many studies of the acquisition of vowel reduction of L2 English have been held, registering that intensity and duration could not be produced in a native-like fashion, that early bilinguals centralize the vowels more than late bilinguals(Rallo Fabra, 2015) and that training improved the perception of the schwa (Lacabex et al., 2008), there is no equal amount of studies that study vowel reduction of L2 Spanish.

To conclude this short section, it has to be argued that the gap in the literature here shown requires a detailed analysis of the phenomena in Interlanguage, as a difference can indeed be perceived in L1 English - L2 Italian vowels.

Chapter 3

The experiment

As previously stated in the introduction, the main focus of this dissertation is the acquisition of vowel reduction in Italian as a L2.
As it has been discussed in the previous chapter, although a distinct centralization in the vowel space in unstressed positions is visible, Italian vowel reduction is a topic that is still subject to some contradictions; for example, whether it is only phonetic or it can be considered phonological or if the neutralization of the mid vowels is a consequence of it. Under these circumstances, the lack of literature of vowel reduction in L2 Italian is not surprising.

The aim of this dissertation is, thus, to start filling this gap. Given the limited time available for the experiment and the actual status of the studies around vowel reduction, some boundaries had to be set in order to work towards valid and reliable results. As acoustical data has already lead to a better understanding of the phenomenon in L1 English and L1 Italian, this study focuses on offering a description of formant values of a vowel in two main conditions, stressed and unstressed and in diverse contexts, in L2 Italian speech by L1 English speakers. Considering that centralization is more evident in peripheral vowels, and that in Italian the effects are mostly visible in the F1 of the low vowel (see section 2.2.4), the vowel /a/ was chosen for the experiment.

The data obtained through the experiment will thus be formant values that will be plotted in the vowel space and confronted to the L1 English and L1 Italian, hoping to give more insight to what happens in the interlanguage and to represent a first step to a complete description of the entire phenomenon.

In this section, the general methodology will be exposed and the choices made will be motivated. Results and discussion will follow in the next chapter.

3.1 Participants

The participants were chosen from a group of 18-20 years old native speakers of English, all of them beginner students of Italian at the same university. None of them have been studying Italian for more than 18 months, and further talk with their professor confirmed their current level of Italian. Moreover, none of the students took specific classes or training on Italian pronunciation.
Of the volunteers, bilinguals (as in not native speakers of other languages in addiction to English) and students with non-English parents (that therefore might have been exposed since a young age to another language) were discarded to avoid any bias in the results; however, students of other languages in addition to Italian were not. This choice was made in order not to limit too much the sample, also considering that most of the students study languages.
In the end, the sample was composed of 2 male and 3 female students, who, before the experiment began, signed a consent form approved by their university. To avoid bias related to a particular focus of the speaker to their pronunciation, the aim of the study was not completely disclosed; however, at the end of the registrations, the participants were informed and did not withdraw.

3.2 The procedure

The participants were invited individually to sit in a recording studio and asked to name in Italian the pictures that were showed to them by the interviewer, which was me.
All the interactions between the participants and the interviewer were strictly in English, in order to avoid any possible imitation of cues in the speech. The interviewer intervened when necessary to help the participant understand what was depicted in the pictures and remember the word without giving cues to its pronunciation. This was achieved encouraging the participants when the word they were trying to say was similar to the target word, and notifying them when the word was dissimilar from the target. For example, instead of "ragazzo" (boy), most participants tended to say "uomo" (man); in this case, the interviewer asked

"and a young man?".

The experiment itself was preceded by a preparatory phase in which I observed and helped beginner Italian students of the university as part of a project; the participants were part of it as well. This phase was useful because it gave a measure of the abilities of the students, which was vital in the preparation of the experiment.

3.3 The task

The task chosen for gathering the needed data was to get the participants to name the object or situation represented in a picture. While it is mostly used in psycholinguistic and neuropsychology studies to calculate effects of priming, response times and so on, there are a few reasons why the picture naming task was chosen as the best method to gather the data.

First, as the chosen participants knew Italian at a beginner level, it would have been difficult to gather the registrations through other means. The preparatory phase showed that the participants felt not able, as a consequence of shyness, hesitancy and fear to make mistakes, to engage in a conversation in Italian even in an informal and friendly setting. As the picture naming task does not require knowledge of grammar and syntax but just of the vocabulary, the students felt more at ease and accepted to participate.
Moreover, the stimuli to be uttered could be chosen in advance as to match both the requirement of the study and their vocabulary knowledge (see section 3.4.1).

While a conversation would have had to be pushed so the same stimuli could be pronounced by all the participants with no guaranteed results, the picture naming task allowed a measure of control over what is said. Although the recordings are not as natural and lack the effects of connected speech, the data that can be extrapolated is homogeneous; moreover, the effect of sentence accent, described by van Bergem (1993, see section 2.2.1), is annulled and thus does not represent a factor of variability.
The picture naming task does therefore allow control of the words said, so that the stimuli can be accurately chosen, and leads to a greater homogeneity since every participant utters the same words in the same context.

As a consequence of the control, the words could be carefully chosen not only to match the participants' knowledge of Italian, but also so that the vowel /a/ may occur in different conditions.

Given the relatively small number of pictures to name (see section 3.4.1), the task takes about five minutes to complete, including potential corrections. While it has the downside of allowing more time to remember the target word compared to connected speech, it helps to avoid fatigue-related issues.
It can be argued that the the attention may be focused on the pronunciation when preparing to name the picture; however, as the participants did not have enough information about the aim of the study and the task itself should suggest more of a "lexical" test (as in testing the participants' vocabulary knowledge), the argument should not hold.

Lastly, as there was no need for a partner to speak with and to direct a conversation, the participants could not imitate the cues in someone else's Italian speech, thus reflecting their own ability.

In conclusion, while the most important reason for choosing the picture naming task was to allow gathering the data from beginner speakers, it also had the merits of permitting control over the speech of the participants and of avoiding fatigue effects. Although it does not completely accurately represent natural speech, it represents that was believed to be the best method in the circumstances previously explained.

Even so, it had the limit of not representing connected, natural speech, which would have been the preferred method of data gathering in the case of more experienced speakers.

3.4 Materials

3.4.1 Stimuli

The stimuli provided were 24 printed pictures of about 10 X 9,5cm relative to 24 words selected beforehand from the participants' Italian manual, according their presume knowledge of the vocabulary.
Moreover, all the words were chosen on the basis of two other features:

1. they had to be three syllables long: as it was said in the previous chapter, vowel reduction occurs, with the exception of function words in connected

Italian	English
àcqua	water
calciatòre	football player
càmera	room
cartolìna	postcard
cellulàre	mobile phone
farmacìa	pharmacy
fontàna	fountain
formàggio	cheese
francobòllo	stamp
lavàgna	blackboard
ospedàle	hospital
panìno	sandwich
panoràma	view
passapòrto	passport
patàta	potato
piànta	plant
quadèrno	notebook
ragàzzo	boy
ristorànte	restaurant
spaghètti	spaghetti
tàvolo	table
teàtro	theatre
università	university
vacànza	vacation

Table 3.1: List of the selected words in Italian with the English translation

speech, in words with two or more syllable. As the average structure of two-syllable Italian words is CVCV, with the stress usually falling on the first vowel, to guarantee both diversity in the stimuli and the participants' familiarity with the words (which would have been less in four-syllable words), three-syllable words were the more fitting choice.

2. all the words had to contain at list one /a/, whether stressed or unstressed. In table 3.1, it is showed where the stress falls in each word.

This set of words has 16 stressed and 26 unstressed /a/. The unstressed /a/ occur in both non-final and final positions, respectively 17 and 9 times. Moreover, they occur next different consonants (stops, fricatives, liquids, nasals and affricates) or vowels.

The images were chosen from different sources from internet free of copyright and modified to fit in a 10 X 9,5cm rectangle (see Appendix A). Before the experiment they where shown to two volunteers in another picture naming task to evaluate their adequateness; if recognizable, they were printed in colour and cut.

3.4.2 Recording

The participants booked with the interviewer individual sessions in the recording studio of the University of Leeds. They were invited to sit in front of a microphone and instructed to name in Italian the pictures shown to them. The sessions lasted around 5 minutes for every participant and were recorded using the software Audacity. The recordings were saved in the .wav format and the eventual too-high and too-low values were fixed using the "Normalize" effect in the software at -3dB (figure 3.1). All the 5 recordings were cut to make about 24 of them, one per word (see figure 3.2); they were then opened in Praat. With the help of a script called "textgrid_create"[1], a grid was created to select and annotate in the spectrogram the vowel /a/ (see figure 3.3).
The vowel was selected from the point where the formants started to where they ended, with some exceptions.

[1]The script, kindly offered by professor Leendert Plug, "goes through a given directory, bringing up each soundfile and for each creating a matching textgrid, automatically saving any changes you make to the textgrid"

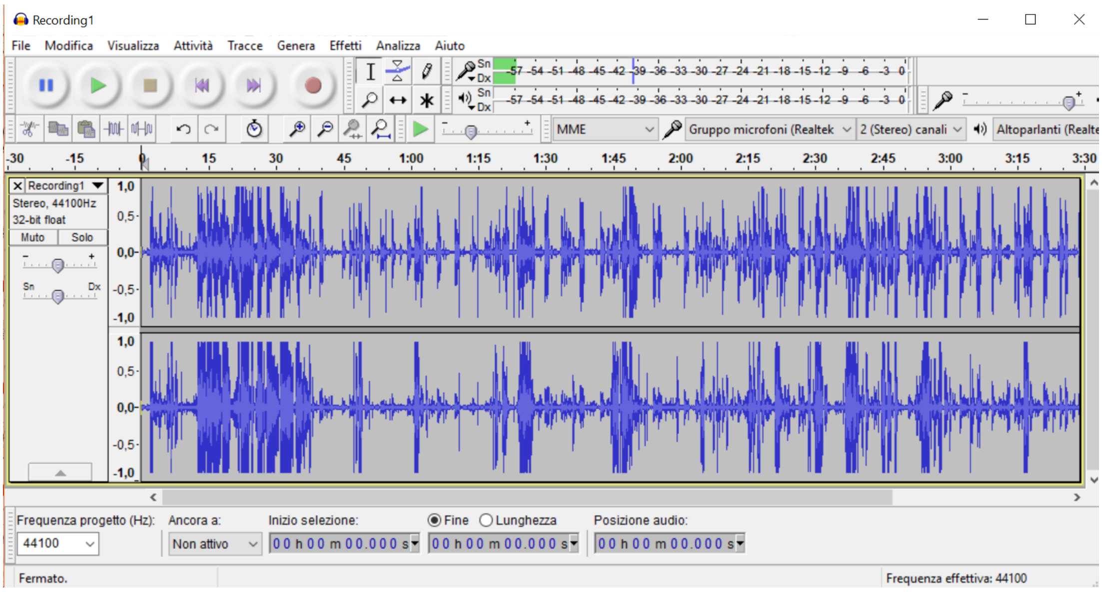


Figure 3.1: The recording before being modified

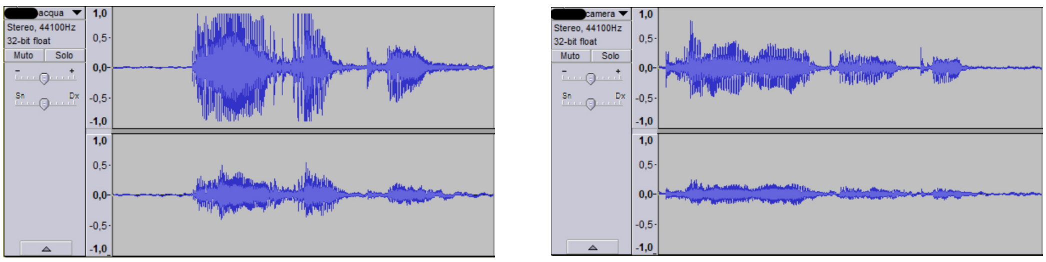


Figure 3.2: An example of individual word recordings

While, in fact, next to stops, fricatives and affricates it was relatively easy to identify the start and end point of the vowel (as visible in figure 3.3), it was difficult to isolate it when it occurred next to liquids and nasals or in diphthongs and monophthongs. While for all of these cases it was chosen to isolate the vowel when it started to sound as an /a/ or something similar, it still presented a difficulty that prevented a valid analysis of the *duration* of the vowel.

With the help of another script called "formant_measure"[2], the values of the F1, F2 and F3 in different points of the annotate /a/ were extracted and inserted in an Excel file. Only the middle values were kept for all the recordings, manually checked and, in some cases, corrected.

[2]This script, written by professor Leendert Plug, "measures formants at given points based on textgrid intervals. It goes through a given directory and writes the results to a .txt file in the same directory".

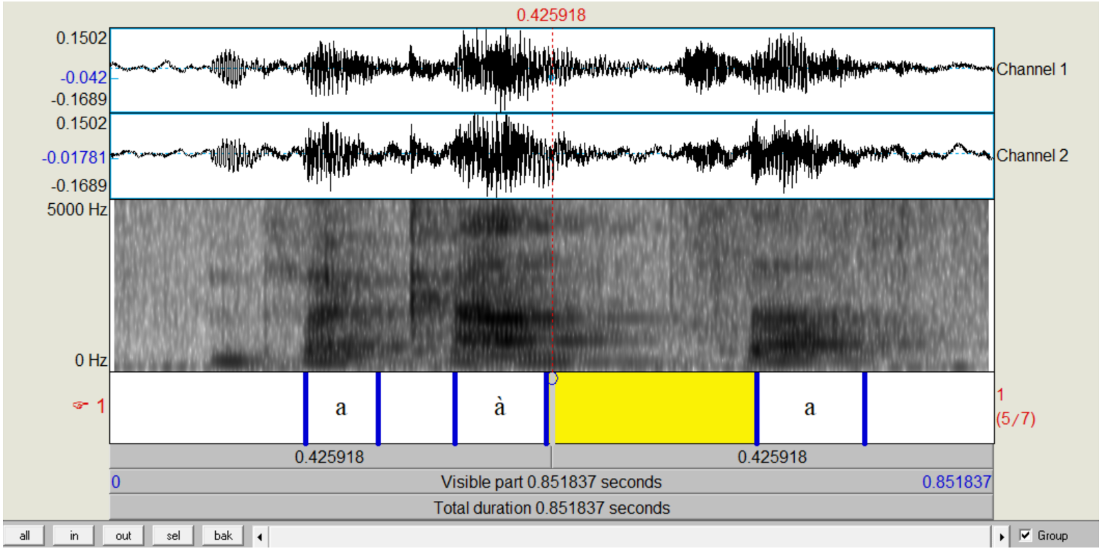


Figure 3.3: The recording in Praat

3.5 Issues

As already anticipated, this kind of study has some issues. Some of them were already explained, such as the impossibility, in the first stage, to extrapolate the data from natural speech and the difficulty of selecting the vowel in particular contexts during the extrapolating stage; other problems however were not.

While every picture was previously evaluated, some of them, especially the one for "panorama" (see appendix A), were not recognised. Moreover, none of the participants could not remember all the words; the interviewer tried in most cases to suggest them in English or reminding them the first letter of the word, but it didn't always work.

Another issue was that in some cases the pronunciation was distinctively not correct, as, for example, the word "ospedale" pronounced as "ospidile" or "ospitale". When a vowel distinct from /a/ or a consonant were mispronounced, as in "ospitale", the error was not acknowledged in the analysis.
When, instead, the vowel /a/ was mispronounced, a perception test was run: if it sounded as a distinct other main vowel, as in the case of "ospidile", it was not included in the analysis.

Lastly, in few cases the word stress was misplaced, as in the case of the word "università"; again, a perception test was run to establish where the accent fell and the data was treated accordingly.

Chapter 4

Results and discussion

4.1 Individual participant data

In this section, the data for each participant will be presented individually; this is to acknowledge intra-speaker variations before giving a more general judgement. In the following tables, all the values in Hertz for the F1 and F2 for each /a/ are summarised. The histograms show the mean values in Hertz for the F1 and F2 for each category, while the black bars summarize graphically the variance (the deviation from the mean). Moreover, all the values are inserted in the vowel space to give a more complete outline of the values with a reference to the English schwa.

With the help of an expert statistician, the t-test was also run for the F1 to asses their significance. As in the literature it is stated that the F2 is variable and generally dependant on the context, it was decided that the test wouldn't be run for those values.

4.1.1 Female speakers

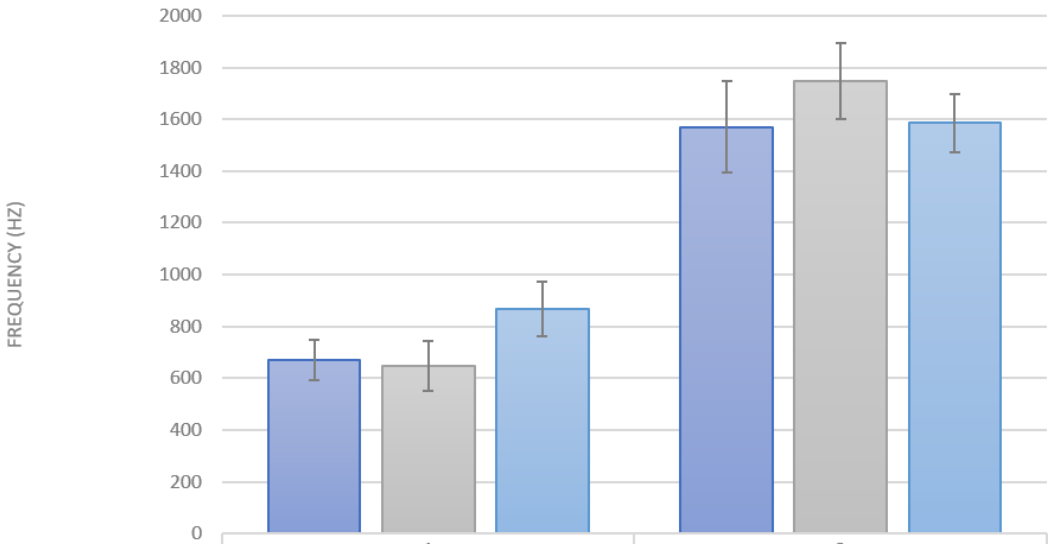
Participant A

The analysis of the F1 and F2 of participant A in Praat yields the following results (see 4.1).

The values for the F1 of all unstressed /a/ are scattered between 564 and 698 Hz. When inserted in the vowel space (see graph 4.2), they however seem to concentrate in the vicinity of the values of the English schwa indicated in section

Non-final unstressed		Final unstressed		Stressed	
F1	F2	F1	F2	F1	F2
680	1243	580	1879	987	1563
676	1581	590	1868	938	1471
698	1592	675	1515	915	1679
846	1336	592	1735	747	1521
564	1449	806	1743	932	1552
687	1625			748	1478
604	1655			985	1572
659	1704			997	1694
694	1760			846	1587
599	1763			806	1693
				804	1347
				687	1782
				964	1629
				788	1627

Table 4.1: Formant frequencies (Hz) of participant A



	1	2
Non-final unstressed	670,7	1570,8
Final unstressed	648,6	1748
Stressed	867,4285714	1585,357143

Figure 4.1: Graph of the average values and displacement of the frequencies of participant A

2.2.3. With a few exceptions, the F1 of the unstressed vowels is sharply distinct from the average F1 values of the stressed /a/, which has an average of 867 Hz (see graph 4.1).

While the difference between unstressed and stressed vowels is without a doubt significant, for unstressed mid and final /a/ the difference in the values is not enough to be significant. The average values are, in fact, respectively 670 and 648 Hz, with similar variance over shared values (see 4.1).

Still in relation to the F1, the words "passaporto", "patata" and "teatro" show the values most divergent from the average.

In the first case, the first /a/ is not reduced, its F1 being 846 Hz. An early explanation lies on the fact that the word "passaporto" has a secondary accent right on the first /a/. As the other word with a secondary accent ("francobollo") was not pronounced by participant A, it wasn't possible to verify this hypothesis on the basis of other data from this participant. However, a comparison to the other participants showed a similar raise of the F1 in comparison to the other words. This seems to support the hypothesis.

As for the word "patata", the final /a/ has an F1 of 806 Hz. A perception

test showed that the word was articulated with particular care; moreover, the participant had showed some uncertainty over the final vowel of the word, and had articulated it with a questioning intonation, as if expecting a confirmation. Thus, it can be argued that was a case of hyperspeech.
Lastly, for the word "teatro", the F1 was of 687 Hz, instead of an expected higher frequency. This might be the result of the effect of the near /e/ or of difficulties in the analysis. As the other participants do not show a similar pattern, it might be the result of difficulties in the analysis; however, there is not enough data to confirm or discard this hypothesis altogether.

As for the F2, while the average values show a sharp divergence, this seems to be the result of the inclusion of some extremely deviant data rather than where the values seem to concentrate (see graph 4.2 and graph 4.1). In other words, the difference in the average values may be the result of a mathematical abstraction that includes some exceptions.
An analysis of these exceptions, that are related to the words "farmacia" (F2 = 1243 Hz), "passaporto" (F2 = 1336 Hz) and "tavolo" (F2 = 1347 Hz), did not suggest any possible explanation not related to an error in the extrapolation of the formant frequencies. Divergent data seem however to be common for all participants.

Another characteristic of the F2 is its variance. The frequencies vary vastly, from a minimum of 1243 to a maximum of 1879 Hz, probably, as suggested in the literature, as a consequence of the different contexts.

Participant B

The F1 of participant B tend to show more regularity in contrast to participant A. While the variance for the non-final unstressed vowel is high, between 597 and 847 Hz, for final unstressed and stressed /a/ the values seem more consistent, not far from the average of, respectively, 800 and 924 Hz (see 4.3).

The non-final unstressed vowels seem to be divided in two main groups, one scattered near the F1 value of the English schwa and one overlapping with final unstressed vowels and almost overlapping with stressed vowels.
This phenomenon may have different explanations:

1. all the unstressed vowels underwent reduction, but in different gradations. The blank space between the two groups may be, in this case, a random

Non-final unstressed		Final unstressed		Stressed	
F1	F2	F1	F2	F1	F2
847	1552	761	1288	876	1530
597	1874	790	1634	923	1559
773	1373	815	1484	975	1330
827	1390	816	1510	937	1562
782	1586	793	1576	902	1629
796	1755	858	1396	870	1431
806	1453	770	1389	932	1421
605	1420			942	1471
636	1829			842	1623
725	1389			967	1521
761	1676			904	1467
604	1789			987	1787
615	1705			934	1525
				945	1472

Table 4.2: Formant frequencies (Hz) of participant B

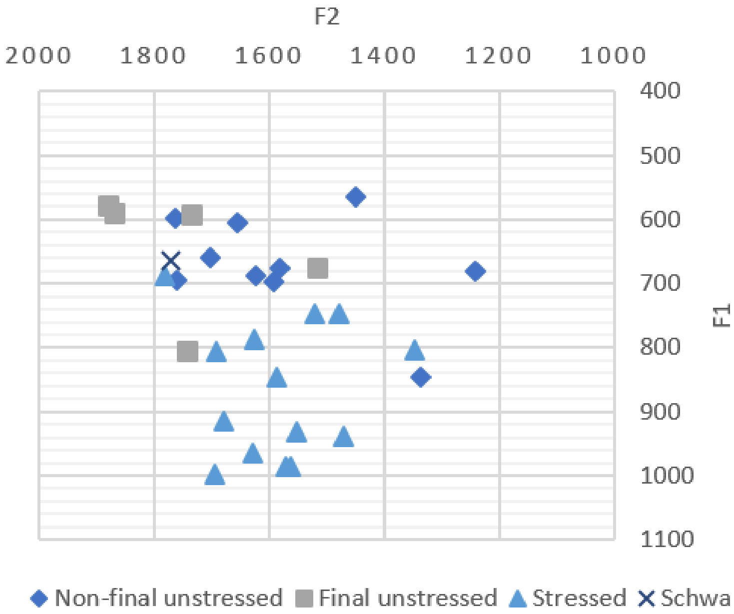


Figure 4.2: /a/ of participant A in the vowel space

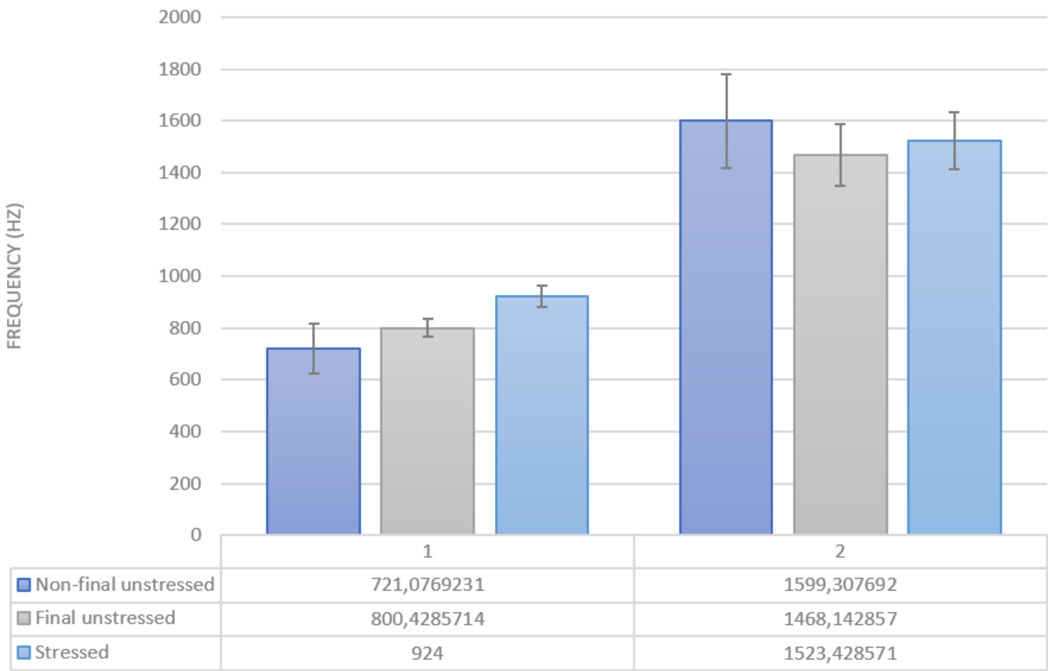


Figure 4.3: Graph of the average values and displacement of the frequencies of participant B

occurrence, that would have been avoided with a bigger set of stimuli;

2. the group nearer the stressed vowels is not reduced. This is supported by the fact that vowel reduction is not a mandatory rule in English, and thus was applied to some vowels and not to others; however, a bigger overlap with the stressed vowels should have been more evident.

No individual data could be found in the literature to support one hypothesis over the other. However, as the statistics demonstrate that there is indeed a significant difference between unstressed and stressed vowels, the second hypothesis may be rejected.
As for middle and final unstressed /a/, the statistics show that the difference, however significant, is minimal.

The F2 behaves similarly to participant A's. Graph 4.4 shows an evident divergence of a few particular /a/, especially for the lower values (which are in the higher part of the vowel space). The 6 vowels who showed F2 values between 1200 and 1400 Hz occurred in the words "farmacia" and "quaderno" in the middle positions and in "acqua", "pianta" and "vacanza" in the final positions. While they were analized as in the case of participant A, no explanation could be found.

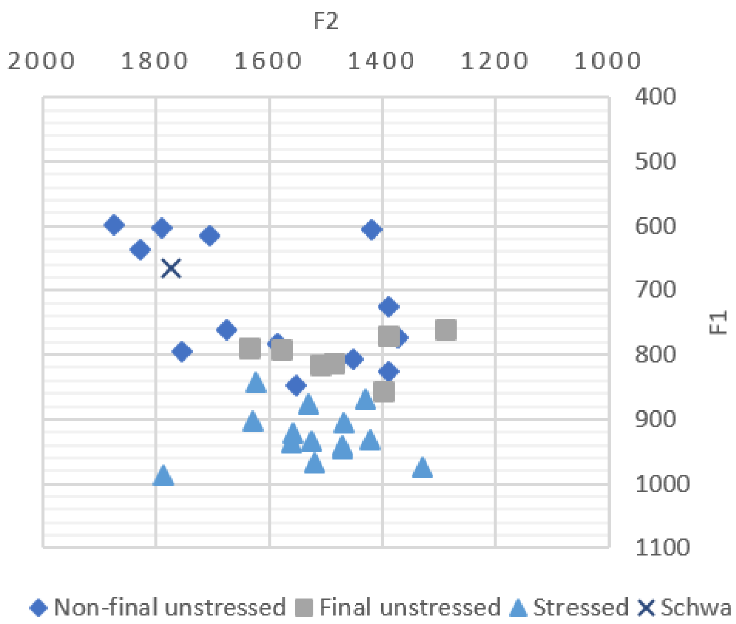


Figure 4.4: /a/ of participant B in the vowel space

Non-final unstressed		Final unstressed		Stressed	
F1	F2	F1	F2	F1	F2
765	1420	675	1524	834	1592
796	1497	642	1468	752	1528
796	1762	734	1881	889	1132
863	1982	704	1835	889	1668
837	1823	742	1658	826	1302
703	1789	796	1899	903	1700
514	1751			781	1951
606	1640			959	1823
553	1784			871	2126
609	2040			878	1645
561	1975			980	1596
				952	1820
				977	1806
				830	1849

Table 4.3: Formant frequencies (Hz) of participant C

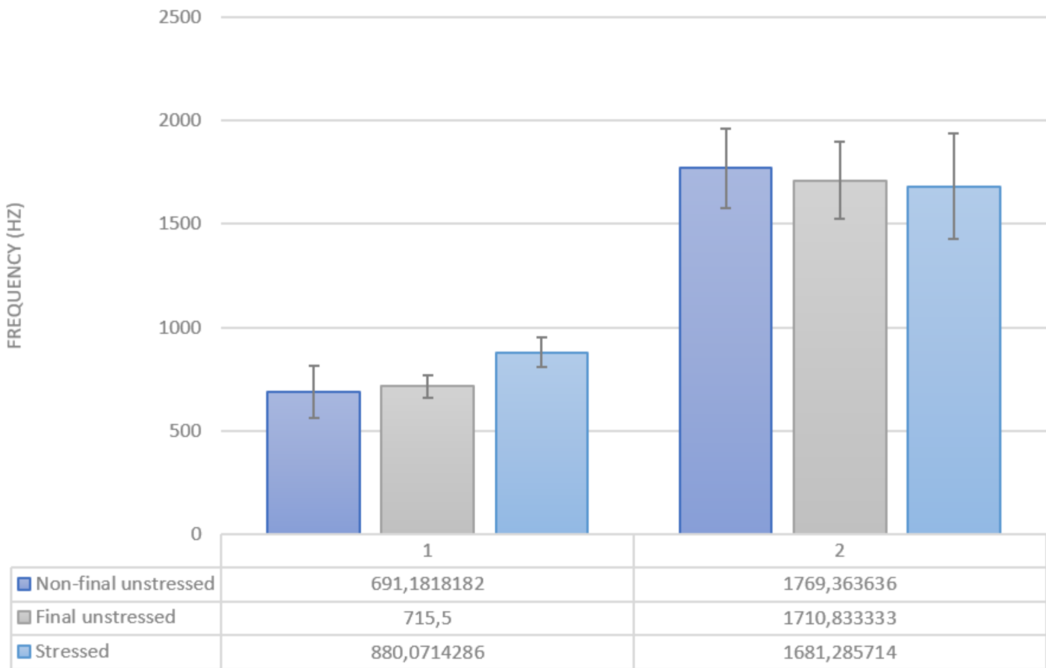


Figure 4.5: Graph of the average values and displacement of the frequencies of participant C

Participant C

It is immediately visible from graph 4.6 (and partly from graph 4.5) the distinct overlap between the unstressed and stressed vowels. The /a/ are mostly scattered in the vowel space, not concentrating in the vicinity of the averages; notably, the F1 of non-final unstressed vowels seem to behave as they do in participant B, dividing in two groups, although less clearly defined.
While the vowel space seems peculiar, the averages of the F1 (see graph 4.5) are in the norm if compared to the other participants and to the graphs shown earlier (figures 2.5 and 2.7). Moreover, the t-test shows a significant difference between unstressed and stressed vowels; again though, there is no significance between middle and final unstressed vowels.

Although the variance, especially for the non-final unstressed vowels, is high, no particular value of the F1 stands out for being isolated and far from the group (see 4.6). Some limit cases are represented by those values on the edge of the group, as the unstressed /a/ in the words "panino" (F1 = 863 Hz) and "passaporto" (first /a/, F1 = 837 Hz) and the stressed /a/ in the words "camera"

(F1 = 752 Hz) and "patata" (F1 = 781 Hz).
With the exception of "passaporto" that, as previously discussed, can be identified as a special case due to its secondary accent, all these cases show values not far from the norm and may be random.

The F2 show an even more pronounced variance, evident in both the graphs 4.5 and 4.6. This variance is certainly related to the divergence that some data shows.
An analysis of the vowels with F2 frequencies lower than 1500 Hz or higher than 2000 Hz showed a common pattern, which is their proximity to nasals and liquids. While in these cases all the conflicting /a/ are next to nasals and liquids, not all the /a/ in this context show the divergence. It is the case, for example, of the word "ristorante": while the /a/ is next to an /r/ and an /n/, its F2 is of 1645 Hz, which is average compared to the 1132 Hz of the F2 of /a/ in the word "cellulare".
Moreover, the divergence pattern in not present in the other participants' data. It is therefore safe to assume that it is due to errors in the analysis itself, considering that, as it was stated in the previous chapter, there were some issues especially in these contexts.

If the conflicting data is removed from the analysis, the pattern of the F2 seems otherwise normal.

4.1.2 Male speakers

As there are some general widely known differences between the frequencies in male speech versus female speech, some changes had to be made. The formant values of the schwa, that is used in the vowel space graph as a reference point, had to be adjusted; additionally, the scale of the axes, as it had to accommodate lower values, was modified.

Participant D

A first look at the formant values of participant D reveals an absence of overlap between non-final unstressed and stressed vowels.

The formants of the stressed /a/ tend to concentrate around the average values with little variance. There is just one atypical case represented by the

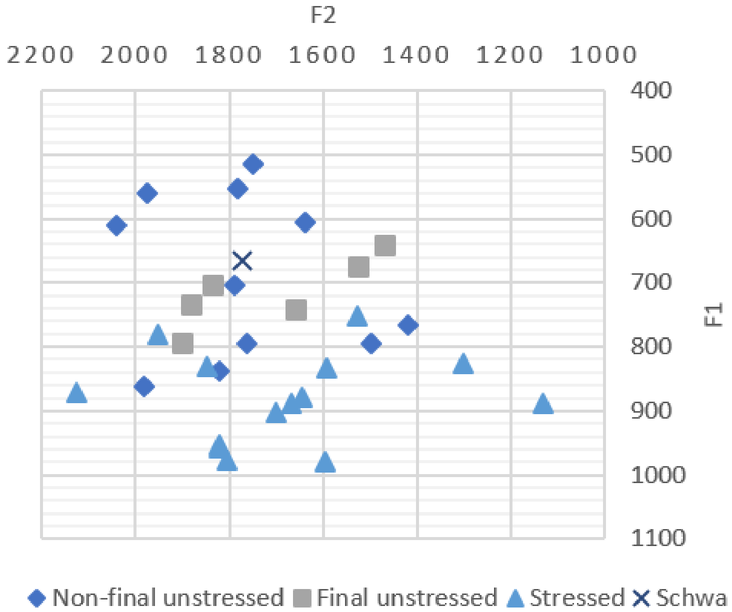


Figure 4.6: /a/ of participant C in the vowel space

word "ristorante" (F1 = 731), that may be due to an error in the analysis, given the proximity of the /r/ and /n/ (similarly as to what happened for participant C), or to a random variation.

The unstressed values show instead more variance, both for F1 and F2.

The limit case, which is not depicted in graph 4.9 due to lack of space, is represented by the word "passaporto". The second /a/ has, in fact, a decidedly low F1 (291 Hz) and a remarkably high F1 (2590 Hz), values that are generally associated with a front high vowel[1].

As the vowel is clearly perceived as an /a/ and not an /i/, it seems clear that the fault is in the analysis itself. From an ulterior analysis of the spectrogram (figure 4.8) it is possible to elaborate some hypothesis:

1. the background noise might have covered the formants of the /a/.

2. the formants of the vowel have been somewhat assimilated by the precedent consonants. This hypothesis is supported by the presence of what appear to be formant lines before what has been indicated as the start of the vowel; however, there is no trace of voicing until after the probable formants end.

[1] Ladefoged and Johnson (2011) indicate the values of an English [i] as 280 Hz for the F1 and 2250 Hz for the F2.

Non-final unstressed		Final unstressed		Stressed	
F1	F2	F1	F2	F1	F2
494	1313	512	1168	640	1365
470	1546	602	1316	640	1263
477	1751	602	1441	571	1363
415	1223	508	1460	601	1510
291	2590			700	1425
540	1285			696	1488
415	1316			664	1472
512	1409			731	1254
490	1571			662	1350
384	1667			664	1551
				633	1316
				607	1466

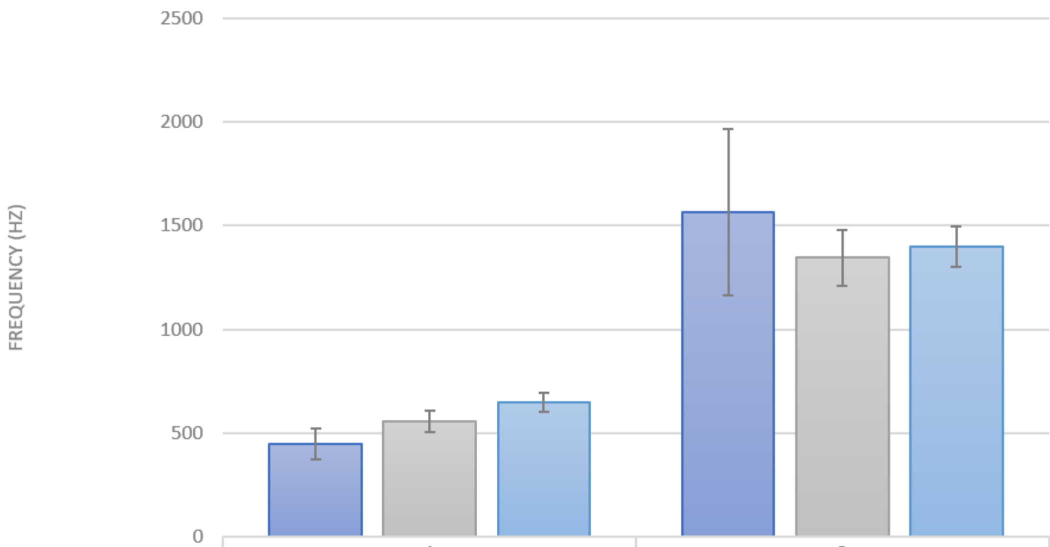
Table 4.4: Formant frequencies (Hz) of participant D

3. there is no vowel. In this case the results of the perception check might not reflect the acoustical nature of the word, but a subsequent manipulation by the brain.

As there is not enough data to support one hypothesis over the other, it was decided to not proceed to further analysis of this case. It must be remembered, however, that the presence of the formants in the dataset alters the average and the variance of the non-final unstressed vowels.

As for the F1, other three cases, although not as conflicting as the previous one, clearly deviate from the average. An already discussed and explained one is represented by the word "passaporto", with an F1 of 415 Hz. The other cases are related to the words "vacanza" (first /a/, F1 = 384 Hz), which also shows a deviant F2 (1667 Hz), and the word "quaderno" (F1 = 415 Hz).
While for "quaderno" the explanation is most probably related to the proximity of the /u/, which might have caused either the lowering of the formant or an error in the analysis, for "vacanza" it may be a random effect on both the F1 and the F2.

One last deviant case involves only the F2 and is represented by the word



	1	2
Non-final unstressed	448,8	1567,1
Final unstressed	556	1346,25
Stressed	650,75	1401,916667

Figure 4.7: Graph of the average values and displacement of the frequencies of participant D

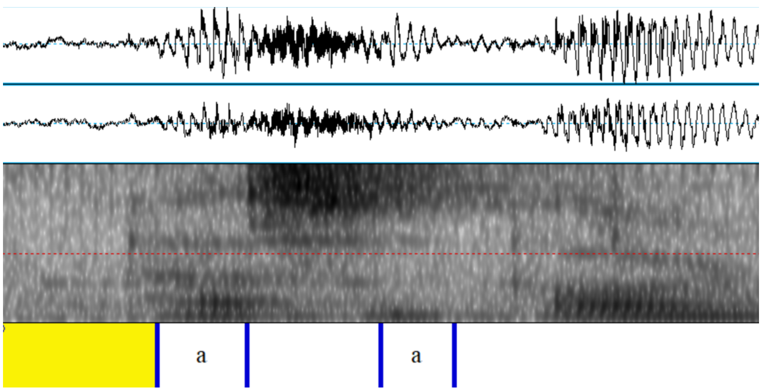


Figure 4.8: Spectrogram of the word "passaporto" as uttered by participant D

"panino" (1751 Hz). As there are no similar deviations caused by the nasal in the data, it seems to be again the consequence of a random effect.

To conclude, the t-test showed a significant difference between unstressed and stressed vowels in the F1. While it does also show a minimal significant difference between middle and final unstressed vowels, as the instances of final vowels are just four, it was decided to not take into consideration this last result.

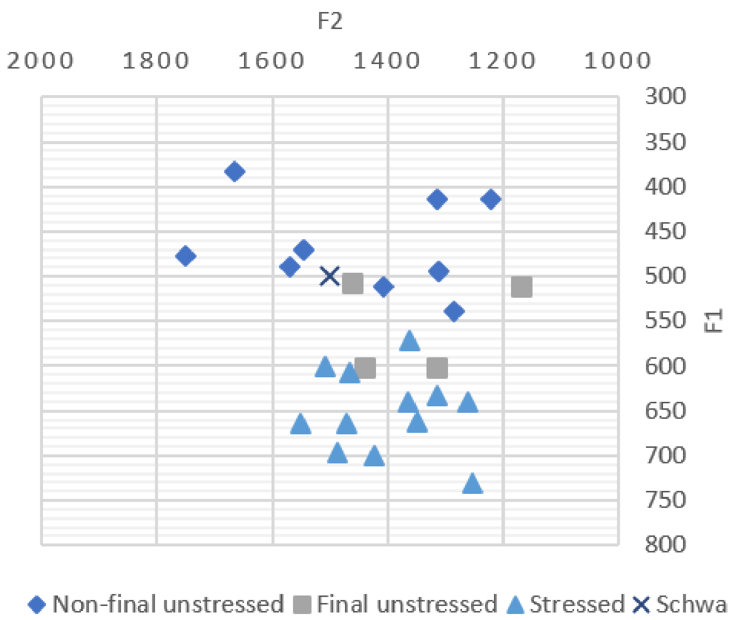


Figure 4.9: /a/ of participant D in the vowel space

Participant E

Participant E's formants immediately appear to be scattered across the vowel space.

As for the F1, a first look at graph 4.10 shows different averages for the variables and a medium level of variance.

While in every group there is at least one or more divergent data, the final unstressed /a/ are the most interesting: they, in fact, seem to be mixed with the other non-final unstressed vowel instead of occupying a small space between them and stressed vowels as it was evident in participant B and C's vowel space (see graph 4.4 and 4.6). The t-test further shows that there is a non-significant difference between the two unstressed groups.

The most divergent values refer to the words "cartolina" (F1 = 664 Hz) and "camera" (F1 = 602 Hz). A perception test run on both worst showed that the former was uttered with great care; it is not therefore surprising that it shares values with stressed vowels. The latter was not pronounced in isolation but in the phrase "camera da letto" (bedroom). As speech is a continuum, it may be argued that the vowel should be treated as a non-final instead of final.

As for the other two groups, while some overlap in the range between 600

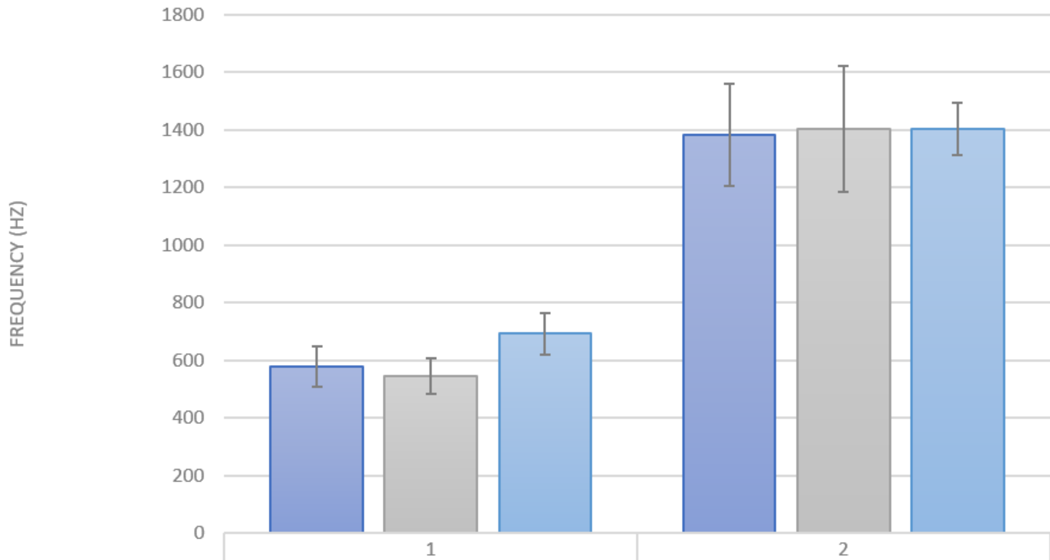
Non-final unstressed		Final unstressed		Stressed	
F1	F2	F1	F2	F1	F2
664	1410	508	985	819	1410
665	1038	602	1541	602	1285
575	1321	664	1432	695	1348
575	1420	478	1631	725	1348
592	1656	544	1254	713	1344
642	1192	492	1654	665	1389
519	1232	566	1334	766	1459
429	1455	516	1399	695	1563
542	1548			650	1343
611	1544			726	1274
553	1403			672	1489
				534	1525
				733	1466

Table 4.5: Formant frequencies (Hz) of participant E

and 650 Hz seems reasonable, considering that male speech is lower than female speech, some values of stressed /a/ appear too low. It is the case of the words "università" (F1 = 534 Hz) and "camera" (F1 = 602 Hz). While the latter word might have been influenced by the insertion into a wider phrase, in the case of "università" the vowel was pronounced as non accented.

Even with conflicting data, the t-test reported that the difference between unstressed and stressed /a/ is significant.

The variance for F2 is particularly high for unstressed vowels.
The most limit case is represented by the word "acqua" (final /a/, F1 = 508 Hz, F2 = 985 Hz) which, for space reasons, could not be represented in graph 4.11. This could be the consequence the effect of the proximity of the /u/, a high, back vowel that, according to Ladefoged and Johnson (2011), has an F1 of about 310 Hz and an F1 of about 870 Hz.
Some other divergent data come from the words "farmacia" (first /a/, F2 = 1038 Hz), for which the analysis was most likely influenced by the near /r/, and "passaporto" (first /a/, F2 = 1192 Hz and second /a/, F2 = 1232 Hz).



	1	2
Non-final unstressed	578,8181818	1383,545455
Final unstressed	546,25	1403,75
Stressed	691,9230769	1403,307692

Figure 4.10: Graph of the average values and displacement of the frequencies of participant E

With the elimination of the divergent cases, which generally can be reasonably explained, for both the F1 and the F2 the vowel space does not seem dissimilar to the other participants'.

4.2 A general overview and discussion

The individual data showed in the previous section are now discussed on a more general level. In the graphs 4.12 and 4.13 the variables were grouped in different circles; as they are meant to be just a guide and not a perfect depiction of the area covered by the values, it was decided not to include the most divergent data.

While this graphs can therefore account for intra-speaker variability, as the frequencies change substantially in relation to the speaker's sex, male and female data were not merged in one exclusive graph.

Lastly, these graphs use the Bark scale. As stated in section 2.2.4, it has the merit of better reflecting, in comparison to the Hertz scale, how the vowel is perceived.

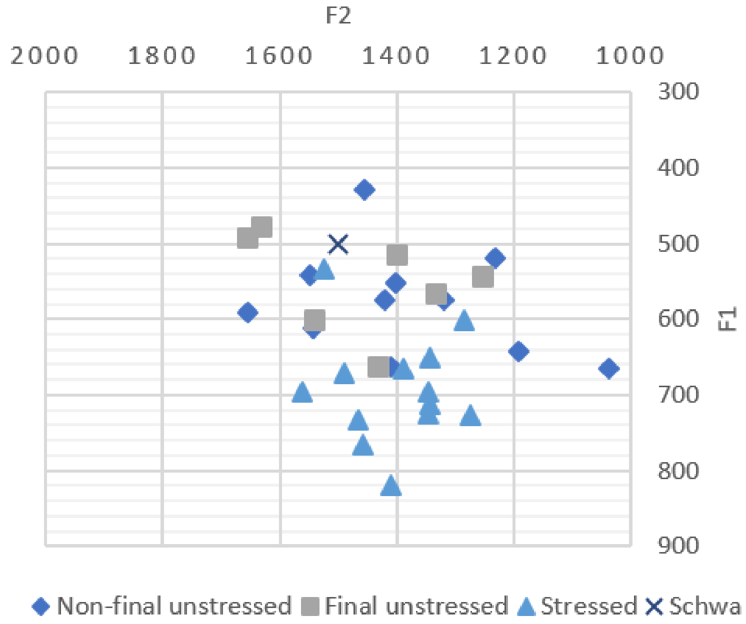


Figure 4.11: /a/ of participant E in the vowel space

In comparison to the figures 2.5 and especially 2.7, which show the extent of vowel reduction respectively in English and Italian, both of the graphs show some peculiar characteristics, which can be summarised as the lack of similarity between the vowel systems.

As it has been previously discussed, there is some overlap both in English and in Italian between stressed and unstressed vowels, which of course changes in relation to the style of speech: particularly in Italian, in hypospeech condition the overlap mostly concerns the different vowels, while in hyperspeech it concerns stressed and unstressed vowel (see figure 2.7).
It is not possible to see the overlap between different vowels in the interlanguage as only the vowel /a/ was analysed; the study is however interested in what happens to the same vowel.
While in both the first languages the overlap, whether in hyperspeech or hypospeech, is minimal, that seems not to be the case for the interlanguage. In the graphs 4.12 and 4.13 the overlap is instead meaningful. Moreover, the significant response of the t-test confirms that vowel reduction indeed occurs.

From this observation a few considerations may be inferred:

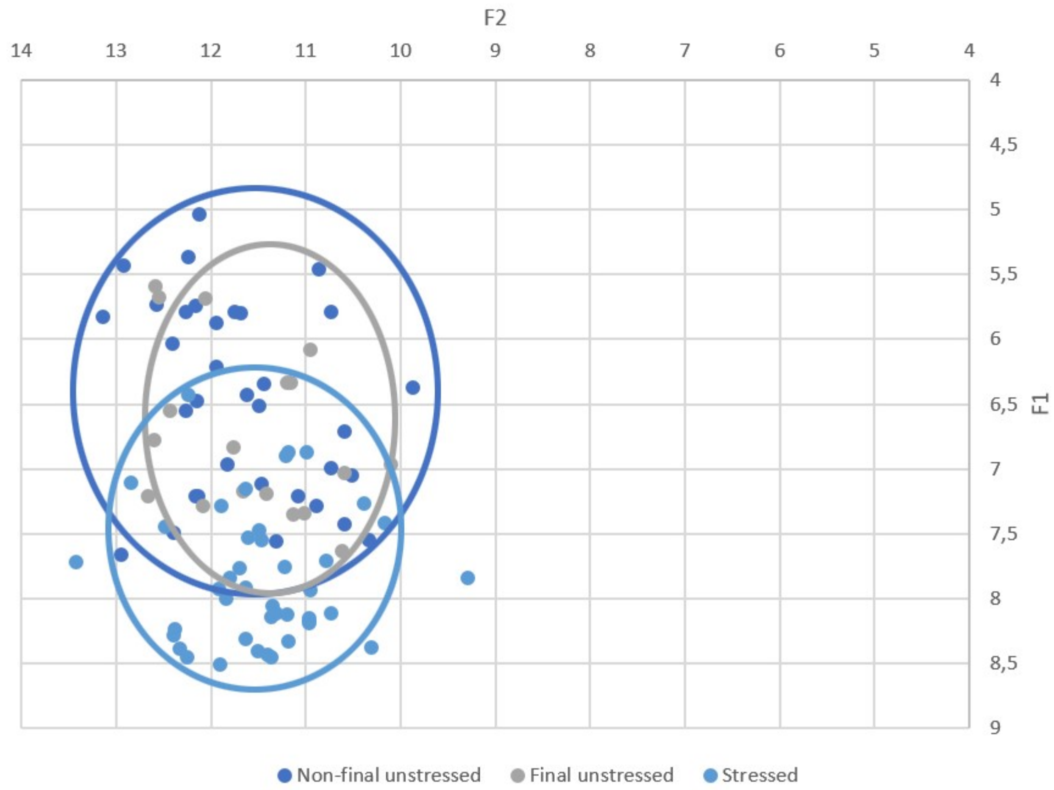


Figure 4.12: The female participants' data in the vowel space (Bark)

- the fault of the discrepancy may lie in the chosen task. As it was already argued in section 3.3, although picture naming has a huge number of merits (it is adequate for beginner learners, allows for more homogeneity of data, avoids fatigue-related issues, doesn't involve a second speaker), it doesn't accurately reflect natural speech.
  As there is no study in the literature that uses this kind of task in SLA, there is no data to support or oppose the first argument. An additional study, involving this task on a control group of native speakers of Italian and native speakers of English, with an additional analysis could offer enough evidence for or against the use picture naming.

- participants may have been hyper aware of their pronunciation. In the case this hypothesis was verified, it would be confirmed that, no matter the speech style and the degree of self-awareness in speaking, phonetic vowel reduction is unavoidable.
  From the analysis of the single speakers, however, some divergent data was

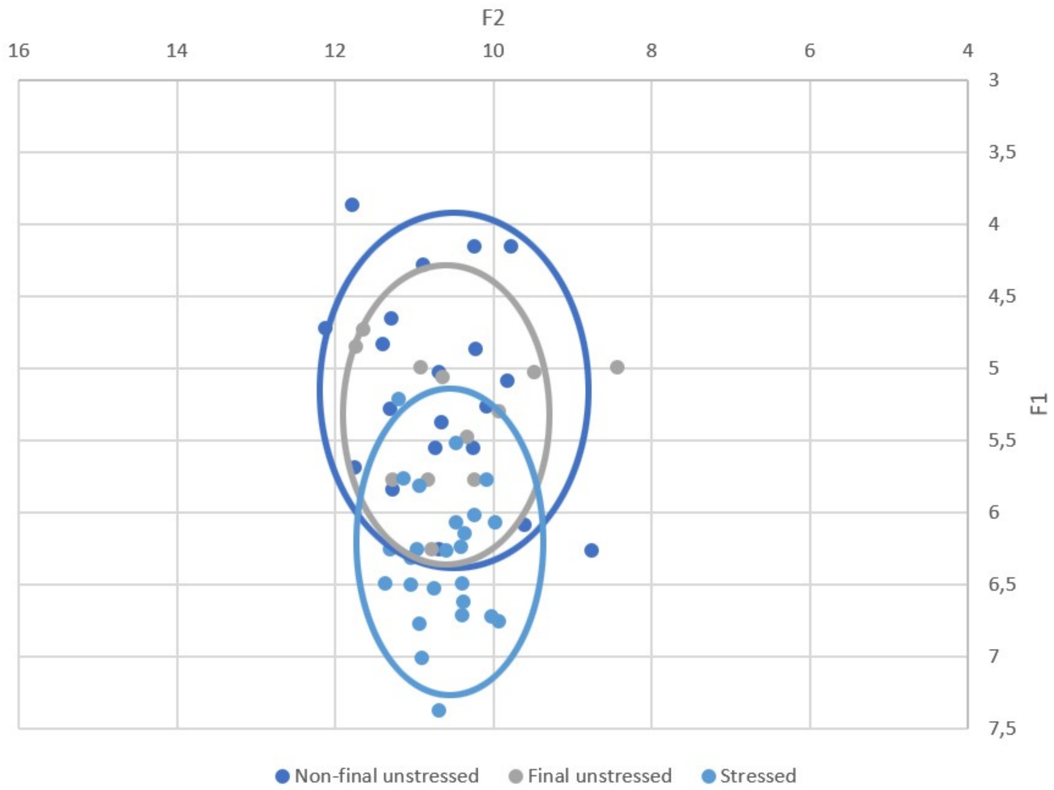


Figure 4.13: The male participants' data in the vowel space (Bark)

justified as being a case of hyperspeech. Assuming that, if the participant were already hyper aware of their utterance, it would not be possible to be even more aware (thus explaining the divergence in the data), there are two possibilities: either the explanation for the divergence is wrong, or the entire hypothesis of hyper awareness is wrong.
In either case, a perception test seems the only plausible may to support one or the other hypothesis.

- the data is also valid for natural speech and reflects a stage of acquisition of the L2 phonology. Although Italian and English have a similar system of vowel reduction, if the data is indeed natural, it would indicate an early phase in the interlanguage in which the rules of both the L1 and the L2 are not yet applied. Even if it were not the case, it would still imply that transfer from the L1 does not always occur, offering a piece of evidence against the Contrastive Analysis Hypothesis.
To further support this hypohesis, a more in-depth study of interlanguage

plus a more naturalistic method of gathering data would be needed.

This considerations still hold true when considering that the data shows a non significant difference between word-final and middle unstressed vowels as opposed to what happens in English and, to a lesser extent, in Italian.

4.3 Conclusions

While the experiment was motivated by a general curiosity as to how vowel reduction occurs in the interlanguage, the results were unexpected. Instead of reproducing the patterns typical of the L1, English, or those of the target language, Italian, vowel reduction in the interlanguage seems be remarkably less pronounced.
To explain these results, a few hypothesis were proposed: the difference may be the result of a bias in the methodology (the picture naming task), of hyper awareness, or of lack of transfer from the L1, thus disproving the Contrastive Analysis Hypothesis. All these hypothesis find a basis not only in the review of the previous studies both on vowel reduction and SLA, but also on a through analysis of the results from each single participant. In addition to the hypothesis, different methodologies were suggested as a pathway to possible future studies.

Of course, the limits of the experiment were not ignored. Problems in the analysis of the formants in the proximity of certain types of consonants not only precluded the possibility of analysing the duration of the vowel in function of its reduction, but also of conducting a through analysis of the variability of the F2 which, as stated in the literature, it the one mostly influenced by the context. In the light of these difficulties, which were not predicted in the preparation stage of the experiment, a less varied context would have been preferred.
The choice of task has been discussed several times, with its pros and cons. A last consideration concerns the possibility of a control group to test its validity. If there weren't time constraints, one would have certainly been assembled and the resulting additional data analysed in order to find a definitive answer.
One last limit was the choice of just one vowel, which was similarly motivated by time constraints.

However so, these results open a new sets of questions that, if answered, may close a gap not only pertaining to the acquisition of vowel reduction, but to the acquisition of a second language's phonetics, as it offers more evidence in support of the Interlanguage theory.

Bibliography

Albano Leoni, F., M. R. Caputo, L. Cerrato, F. Cutugno, P. Maturi, and R. Savy (1995). Il vocalismo dell'italiano : analisi di un campione televisivo. *Studi Italiani di Linguistica Teorica e Applicata 24*, 405–411.

Altman, G. and D. Carted (1989). Lexical stress and lexical discriminability: Stressed syllables are more informative, but why? *Computer Speech & Language 3*(3), 265–275.

Barnes, J. (2002). *Positional neutralization: a phonologization approach to typological patterns.* Doctoral dissertation, UCB.

Barnes, J. (2006). *Strenght and Weakness at the Interface.* Berlin: Mouton the Gruyter.

Bertinetto, P. (1981). *Strutture prosodiche dell'italiano.* Firenze: Accademia della Crusca.

Best, C. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies (Ed.), *Developmental neurocognition: Speech and face processing in the first year of life*, pp. 289–304. Dordrecht: Kluwer.

Best, C. (1994). The emergence of native-language phonological influence in infants: A perceptual assimilation model. In H. Nussbaum, J. Goodman, and C. Howard (Eds.), *The transition from speech to spoken words: The development of speech perception*, pp. 167–224. Cambridge, Mass: MIT.

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

Bybee, J. (2011). *Phonology and Language Use.* Cambridge: Cambridge University Press.

Cobb, K. and M. Simonet (2015). Adult second language learning of spanish vowels. *Hispania 98*(1), 47–60.

Collins, B. and I. M. Mees (2003). *The Phonetics of English and Dutch* (5 ed.). Leiden: Brill.

Cook, V. J. and M. Newson (1996). *La Grammatica universale : introduzione a Chomsky* (2 ed.). Bologna: Il Mulino.

Corder, S. P. (1971). Idiosyncratic dialects and error analysis. *International Review of Applied Linguistics 9*(1-4), 149–159.

Crosswhite, K. M. (2004). Vowel reduction. In B. Hayes, R. Kirchner, and D. Steriade (Eds.), *Phonetically-based phonology*, pp. 191–231. Cambridge: Cambridge University Press.

De Mauro, T., F. Mancini, M. Vedovelli, and M. Voghera (1993). *Lessico di frequenza dell'italiano parlato.* Milano: Etaslibri.

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning 27*, 315–330.

Eckman, F. R. (1991). The structural conformity hypothesis and the acquisition of consonant clusters in the interlanguage of esl learners. *Studies in Second Language Acquisition 13*, 23–41.

Eckman, F. R. (2004). From phonemic differences to constraint rankings: Research on second language phonology. *Studies in Second Language Acquisition 26*(4), 513–549.

Farnetani, E. and D. Recasens (2010). Coarticulation and connected speech processes. In . Hardcastle, W. J., J. Laver, and F. Gibbon (Eds.), *The Handbook of phonetic sciences* (2 ed.)., Chapter 9, pp. 316–352. Malden, MA: Wiley-Blackwell.

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

Flege, J. E. and I. R. A. MacKay (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition 26*(1), 1–34.

Flemming, E. (2004). Contrast and perceptual distinctiveness. In B. Hayes, R. Kirchner, and D. Steriade (Eds.), *Phonetically-based phonology*, pp. 232–276. Cambridge: Cambridge University Press.

Flemming, E. (2009). The phonetics of shwa vowels. In D. Minkova (Ed.), *Phonological weakness in English*, pp. 78–98. London: Palgrave Macmillan.

Flemming, E. and S. Johnson (2007). Rosa's roses: reduced vowels in American English. *Journal of the International Phonetic Association 37*(1), 83–96.

Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical Society of America 90*(4), 1816–27.

Giacalone Ramat, A., M. Chini, and C. Andorno (2014). Italiano come L2. http://www.academia.edu/8258813/Italiano_L2 [Online; Accessed May 2017].

Giavazzi, M. (2010). Martin Krämer (2009) : The phonology of Italian. (The phonology of the world's languages.) : Oxford: Oxford university press. *Phonology 27*(2), 332–341.

Greenberg, J. (1976). *Language universals*. The Hague: Mouton the Gruyter.

Gundel, J., K. Houlihan, and G. Sanders (1986). Markedness distribution in phonology and syntax. In F. R. Eckman, E. Moravcsik, and J. Wirth (Eds.), *Markedness*, pp. 107–138. New York: Plenum.

Hancin-Bhatt, B. (1994). Segment transfer: A consequence of a dynamic system. *Second Language Research 10*, 241–269.

Harrington, J. (2010). Acoustic phonetics. In . Hardcastle, W. J., J. Laver, and F. Gibbon (Eds.), *The Handbook of phonetic sciences* (2 ed.)., Chapter 3, pp. 81–129. Malden, MA: Wiley-Blackwell.

Hart, J. T., R. Collier, and A. Cohen (1990). *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody (Cambridge Studies in Speech Science and Communication)*. Cambridge: Cambridge University Press.

Hay, J. F., M. Sato, A. E. Coren, C. L. Moran, and R. L. Diehl (2006). Enhanced contrast for vowels in utterance focus: A cross-language study. *The Journal of the Acoustical Society of America 119*(5), 3022–3033.

Kager, R. (2007). Feet and metrical stress. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology*, pp. 195–228. Cambridge: Cambridge University Press.

Kingston, J. (2007). The phonetics-phonoogy interface. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology*, pp. 401–434. Cambridge: Cambridge University Press.

Koopmans-van Beinum, F. (1983). Systematics in vowel systems. In M. van den Broecke, V. van Heuven, and W. Zonneveld (Eds.), *Sound Structures*, pp. 159–171. Dordrecht: FORIS Publications.

Krämer, M. (2009). *The phonology of Italian. (The Phonology of the World's Languages.)*. Oxford: Oxford University Press.

Lacabex, E., M. Lecumberri, and M. Cooke (2008). Perception of English vowel reduction by trained Spanish speakers. In A. Rauber, M. Watkins, and B. Baptista (Eds.), *New Sounds 2007 : Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech*, Federal University of Santa Catarina, pp. 294–299. Florianópolis, Brazil.

Ladefoged, P. and K. Johnson (2011). *A Course in Phonetics* (6 ed.). Boston, MA: Cengage Learning.

Lado, R. (1957). *Linguistics across cultures.* Ann Arbor: The University of Michigan Press.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, MA: MIT.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America 35*(11), 1773–1781.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle and A. Marchal (Eds.), *Speech Production and Speech Modeling*, pp. 403–439. Netherlands: Kluwer Academic.

Lindblom, B. (1996). Role of articulation in speech perception: Clues from production. *The Journal of the Acoustical Society of America 99*(3), 1683–1692.

Maiden, M. (1995). Vowel systems. In M. Maiden and M. Parry (Eds.), *The dialects of Italy*, pp. 7–14. London: Routledge.

Major, R. C. (1987). A model for interlanguage phonology. In G. Ioup and S. Weinberger (Eds.), *Interlanguage phonology*, pp. 101–124. Cambridge: Newbury House Publishers.

Major, R. C. (1998). Interlanguage phonetics and phonology : an introduction. *Studies in Second Language Acquisition 20*(2), 131–137.

Major, R. C. (2001). *Foreign accent*. Amsterdam: Benjamins.

Mazzola, M. L. (1976). *Proto-Romance and Sicilian*. Lisse: Peter de Ridder.

Moon, S. and B. Lindblom (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical Society of America 96*(1), 40–55.

Mori, L. (2007). *Fonetica dell'italiano L2*. Pisa: Carocci editore.

Nadeu, M. (2014, sep). Stress- and speech rate-induced vowel quality variation in Catalan and Spanish. *Journal of Phonetics 46*, 1–22.

Nemser, W. (1971). Approximative systems of foreign language learners. *International Review of Applied Linguistics 9*(1-4), 155–123.

Rallo Fabra, L. (2015). Can nonnative speakers reduce english vowels in a native-like fashion? Evidence from L1-Spanish L2-English bilinguals. *Phonetica 72*(2-3), 162–181.

Roach, P. (2009). *English Phonetics and Phonology: A Practical Course* (4 ed.). Cambridge: Cambridge University Press.

Savy, R. (1999). Riduzioni foniche nella morfologia del sintagma nominale nel parlato spontaneo: indagine quantitativa e aspetti strutturali. In P. Benincà, A. Mioni, and L. Vanelli (Eds.), *Fonologia e morfologia dell'italiano e dei dialetti d'Italia : atti del XXXI Convegno della Società di Linguistica Italiana : Padova, 25-27 settembre 1997*, Roma, pp. 201–221. Bulzoni.

Savy, R., G. Clemente, and M. Lo Prejato (2005). Per una caratterizzazione e misura della riduzione vocalica in italiano. In P. Cosi (Ed.), *Misura dei parametri : atti del 1 Convegno Nazionale dell'Associazione Italiana di Scienze della Voce : Padova, 2-4 dicembre 2004*, Brescia, pp. 135–160. EDK Editore.

Savy, R. and F. Cutugno (1997). Ipoarticolazione, riduzione vocalica, centralizzazione: come interagiscono nella variazione diafasica. In F. Cutugno (Ed.), *Fonetica e fonologia degli stili dell'italiano parlato : atti delle 7. Giornate di studio del Gruppo di fonetica sperimentale A.I.A. : Napoli, 14-15 novembre 1996*, Roma, pp. 177–194. Esagrafica.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics 10*(1-4), 209–231.

Stevens, K. (1998). *Acoustic phonetics*. Cambridge, Mass: MIT.

Tarone, E. (1988). *Variation in Interlanguage*. London: Edward Arnold.

van Bergem, D. R. (1993). Acoustic vowel reduction as a function of sentence accent, word stress, and word class. *Speech Communication 12*(1), 1–23.

van Bergem, D. R. (1995). *Acoustic and lexical vowel reduction*. Amsterdam: IFOTT.

Watt, D., A. Fabricius, and T. Kendall (2010). More on vowels: plotting and normalization. In M. Di Paolo and Yaeger-Dror (Eds.), *Sociophonetics : a student's guide*, pp. 107–118. New York: Routledge.

Yavas, M. (2011). *Applied English Phonology* (2 ed.). Oxford: Wiley-Blackwell.

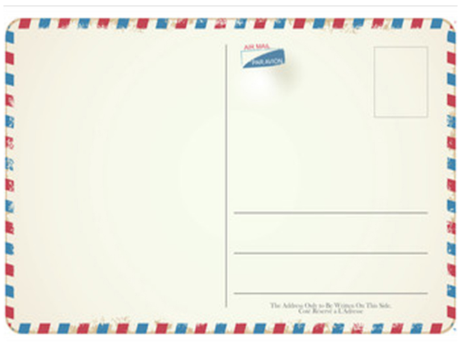
Appendix A

Images

(a) Acqua

(b) Calciatore

(c) Camera



(d) Cartolina

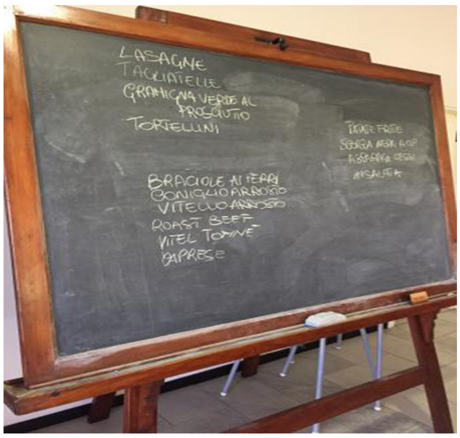
(e) Cellulare

(f) Farmacia

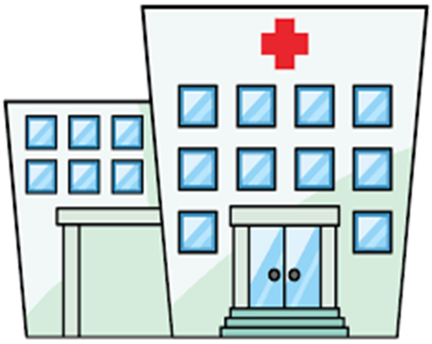
(g) Fontana

(h) Formaggio

(i) Francobollo



(j) Lavagna

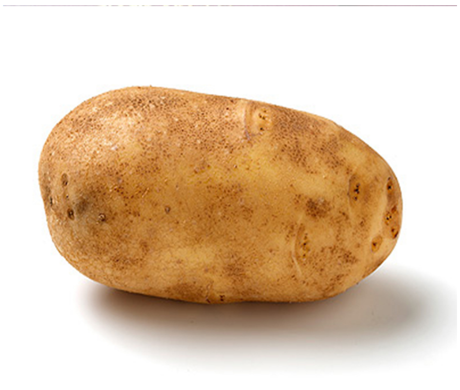


(k) Ospedale

(l) Panino

(m) Panorama

(n) Passaporto



(o) Patata

(p) Pianta

(q) Quaderno

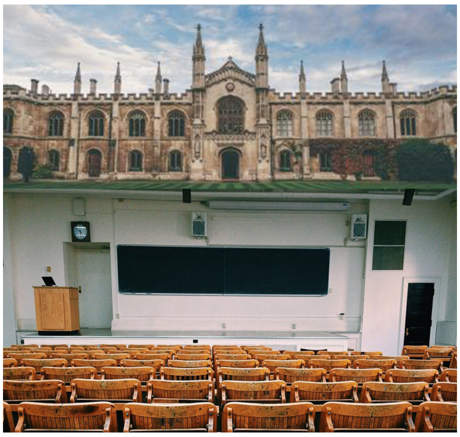
(r) Ragazzo

(s) Ristorante

(t) Spaghetti

(u) Tavolo

(v) Teatro



(w) Università

(x) Vacanza

Figure A.1: Images shown to the participants