A Phonetic Variationist Study on Chilean Speakers of English as a Foreign Language

Paulina Beatriz Subiabre Ubilla

Thesis submitted in fulfillment of the requirements for the
Degree of Doctor of Philosophy
School of Critical Studies
University of Glasgow
2015

April, 2015

ABSTRACT

Variationist research in the Labovian paradigm has traditionally looked at the structured heterogeneity found in first language (L1) speech. More recently this quantitative methodology has been applied to speakers acquiring a second language (L2), usually in immigrant settings. This research has shown that alongside well documented L2 acquisition processes, sociolinguistic patterns are also found, just as in native speech. This dissertation examines the speech of native speakers of Spanish acquiring English in Chile, extending traditional quantitative methodology to L2 contexts, specifically to English as a foreign language (EFL) situations.

I examine the variation of four phonetic variables: voiceless alveolar fricative (ʃ), voiceless alveolar affricate (ʧ), and postvocalic (r), which range from stigmatised to prestigious in both Spanish and English; and voiced dental fricative (ð), which has been extensively documented in English, mainly constrained by linguistic factors. Through the analysis of the speech of eighteen university students, I seek to test, firstly, whether the patterns of variation characteristic of Chilean Spanish are transferred to English and secondly, whether the variation exhibited by native speakers of English is replicated in EFL contexts.

The results suggest that: (1) the expected transfer of patterns from Chilean Spanish to English does not occur for the variables (ʃ) and (ʧ), and (2) the patterns found in non-native speech in EFL contexts replicates the patterns found in native speakers of English for the variables voiced dental fricative (ð) and postvocalic (r). Amongst the social factors considered, the effect of social class is shown to contribute to the variation of postvocalic (r) and (ʃ), as years of instruction in English did to the variation of (ʃ); in relation to the contribution of internal factors, it is found that phonetic environment and position have an effect on the varying use of (ʃ) and (ð). As predicted for (ð), the effect of purely linguistic factors is confirmed. Thus this study demonstrates that the notion of structured heterogeneity can be extended to contexts of EFL, especially in relation to the effect of internal constraints.

ACKNOWLEDGEMENTS

I would like to thank all the people who have helped me through these hard but exciting and fruitful (at times) years:

In first place, my first supervisor Dr Jennifer Smith, for her enthusiastic guidance in the whole process; she showed me that PhD life was not only transcribing, coding, analysing, writing, answering infinite "why's" and "relevance?'s", but also enjoying my stay in Glasgow in different, and always fun, ways; at-all-times concern will always be remembered and, as she once said, "you'd do the same for someone else" (or something like that) – endless thanks, Jen!

To my second supervisor, Prof Jeremy J. Smith, for his great advice and enthusiasm to hear about my progress; his love for linguistic matters was always stimulating and contagious. Also thanks to my third supervisor, Dr Jane Stuart-Smith, for looking after me when my first supervisor could not be present, and for her invaluable advice on phonetics. A thousand thanks to Dr Mercedes Durham for her help in all aspects related to the coding and analysis of my data, and for always promptly and patiently answering my emails.

To the Chilean Funding Programme *Beca Presidente de la República* (later replaced by the programme *Becas Chile*) for funding both my studies and stay in Glasgow; without the scholarship, my research would have been impossible to complete.

Many, many thanks to the whole English Language department staff, for their support and friendship during the four years I lived in Glasgow; special thanks to the cognitive semantics reading group: Prof Christian J Kay, Prof Carole Hough, Dr Wendy Anderson, Dr Fraser Dallachy, Dr Marc Alexander, and Dr Imogen Marcus; it was always fun! Thanks also go to Dr Rachel Smith, Ms Alison Bennett, and Ms Helen Gilday, for their lovely support and help. To the International Student office, especially to Mrs Gillian Gillespie and the International Student Family Network staff, thanks for always welcoming us.

A thousand thanks to my lovely and true friends in the project room (aka the project-room girls): Dr Ellen Bramwell, Dr Johanna Green, Dr Sophie Holmes, and Dr Felicity Maxwell; thanks for all lovely meals, lots of chocolate, biscuits, candy, drinks, the

million cups of tea, and the hugs. You were essential in my life in Glasgow and I've missed you all loads; thanks for always welcoming, helping and supporting me – even from the distance. By the way, I still owe you a couple of postcards!

I would also like to thank my friends in the Latin American community, inside and outside campus: Tania, Alan, Flavia, Carlos, Cristina, Ignacio, Esperanza, Miguel, Teresa, Lucila, Hernan and Jo; you are great friends and excellent cooks. Thanks for the support, help and friendship; you helped making Glasgow our home while we were here.

Thanks dear Michael Lucas (RIP) for your support and friendship; I still have that terrible poem about Glasgow you gave me before leaving Chile, what a laugh it was! Thank you dear Maritza for showing me the "dark side of language" and always believing in me, you've been my inspiration in so many different ways. Also, thanks to the eighteen informants who volunteered to participate in the interviews, your help and patience through the recording process was essential to my investigation. T

hanks to the English teaching programme at Pontificia Universidad Católica de Valparaiso, for asking whether help was needed and for providing it when it was: thanks Carmen, Giovanna, Millaray and Monica; to my anonymous assistants, and especially to my bosses, Eugenia Alfonso, for trusting in me, and Claudia Parra, for the opportunity to work with you and for being by my side as a friend.

Finally, infinite thanks to all my family: first, to my best friend (coincidentally, my husband), Cristobal, for following me in this crazy adventure, for pushing me to finish this, for always believing that I could do it, for putting off his dreams and patiently waiting, and for simply being by my side: I love you. To my brother, Oscar, for everything he has done for me: supporting me, reading early drafts of my work, recommending me the best music for writing up, and babysitting (I love you, too). To my mum, for insistently asking me about my thesis completion and for telling me that I HAD to finish; I might have got over anxious a few times and answered badly, but your nagging was SO necessary – I love you endlessly. To my cousin, Kira, who helped me a lot! To my proofreaders (Kay and Robyn) and my partner in procrastination, Pablo, who helped me with some tedious graphs and lots of nonsensical conversations.

Finally, to my wee sunshine, Vicente: meeting you has been the most wonderful experience I've lived so far. Since you were born, waking up every day and seeing your smile has changed my life; everything I am now is because of you – you're the purest love of my life.

Addendum

During the last months of writing, I would like to thank my *gringuita,* Renata, the final push to this gigantic task. My little baby girl, I wish I had more time to give you during these last months – I love you.

DECLARATION

I declare that this dissertation has been written by me and that the work contained is my own, except where it has been clearly and fully referenced in the text. This original work has not been submitted in any other faculty for another higher degree.

Paulina Beatriz Subiabre Ubilla

Dedication

To my dad, as promised.

I, still, miss you loads.

LIST OF FIGURES AND MAPS

LIST OF TABLES

LIST OF ABBREVIATIONS

NNS	=	Non-native speaker(s)
NS	=	Native speaker(s)
L1	=	First language
L2	=	Second language
ESL	=	English as a second language
EFL	=	English as a foreign language
SLA	=	Second language acquisition
FLA	=	Foreign language acquisition
OD	=	Overall distribution
RP	=	Received pronunciation
SP	=	Spanish
GA	=	General American
SEG	=	Socioeconomic group
YIE	=	Years of instruction in English
PPhE	=	Preceding phonetic environment
FPhE	=	Following phonetic environment
PIW	=	Position in the word
PIT	=	Position in turn
LI	=	Lexical item
V	=	Vowel (e.g. "V_V", means intervocalic position, in which "_" is the feature under study)
C	=	Consonant (e.g. "C_" means that the feature under study is preceded by a consonant)
#	=	Pause (e.g. "#_" means that the feature under study is preceded by a pause)

Contents

CHAPTER 1 - INTRODUCTION

1.1 The study of variation

Since Labov's 1963 report "The social motivation of a sound change", presented on the island of Martha's Vineyard, off Cape Cod, variationist research has traditionally looked at the structured heterogeneity found in L1 speech. Labov's works have since been replicated in numerous speech communities, such as Norwich (Trudgill, 1974), Anniston, Alabama (Feagin, 1979), Toronto (Chambers, 1980), Belfast (Milroy, 1980), Detroit (Wolfram, 1969), Glasgow (Macaulay, 1977; Stuart-Smith, 1999), Buckie (Smith and Tagliamonte, 1998; Smith, 2001), to name a few. The study of sociolinguistic variation has extensively covered the different dialects of English (for a complete account on these, see e.g. Trudgill, 1988a, 1990; Hughes and Trudgill, 1979; Labov, 1969, 1972a; and Labov *et al.* 2006), defining patterns of variation for phonetic, morphosyntactic, and discursive units, and for diverse social and linguistic factors, as in (1) for the study of presence or absence of final coronal stop (taken from Guy, 2007: 7) or (2) for the study of a) presence or b) absence of postvocalic (r) (taken from Labov, 1991 [1972a]: 50):

(1) east side > eas'side
(2) a) casual: fourth floor
    b) emphatic: *fourth floor*

Variationist research is founded on the premise that the variation occurring in all natural languages is the result of the interaction of social and internal linguistic factors that tend to present an "orderly heterogeneity" (Weinreich *et al.*, 1968), and not as a consequence of free, random variation.

Second language (L2) acquisition – although not the focus of this research – has been examined in numerous studies since the late 1960s (see e.g. Corder, 1967; Ellis, 1985a, MacLaughlin, 1987, to name a few). It has also contributed to the study of variation, as it has been argued that interlanguages (IL) show similar behaviour to first languages (L1), and, therefore, they are variable (Adjemian, 1976; Tarone, 1979). More specifically, the quantitative methodology has been applied to studies in L2, most commonly in immigrant settings, in which the varying speech of non-native

speakers (NNS)[1] has also shown distinctive variation patterns (see e.g. Dickerson, 1975; Adamson and Kovac, 1981; Tarone, 1985; Noriega, 2004). During recent years, increased attention has been paid to the study of variation of NNSs in instructional settings (e.g. Schmidt, 1987; Regan, 1996; Major, 1996; Major, 2004; Durham, 2006). In Chilean learners of English as a foreign language, this variation also seems to be present, as shown in the examples (3) – (6), taken from my own data.

(3) (ʃ)[2] > [tʃ]

"I try to just *short* [tʃoɹt] objectives, reachable ones…"

(Paola, first year female speaker)

(4) (tʃ) > [ᵗʃ]; (tʃ) > [ᵗʃ]

"I would like to, to *ch*eat [ᵗʃ], to *teach* [tʃ] at a university."

(Bernardo, third year male speaker)

(5) (ð) > [d]

"I have very good experiences. No, but I'm afraid of, to come back *there* [deɹ] and to be a 'huasa'."

(Paola, first year female speaker)

(6) Postvocalic (r)

"I think *there* [ðeɹ] *are* [ɑːɻ] beautiful landscapes. "

(Consuelo, third year female speaker)

Reserch into examples like these has allowed scholars to question whether variation is also conditioned by the types of constraints documented for native speakers, or if there are different processes at work.

This study on NSs of Chilean Spanish, learners of English as a foreign language (EFL), analyses the speech of eighteen students from the English Teaching Programme at Universidad del Bío-Bío, Chile, at different stages of their English language learning

[1] It is necessary to make a distinction between native speakers' (NSs) and non-native speakers' (NNSs) variation; in the second group, NNSs, two contexts are considered: immigration/immersion and – specifically for this research – English as a foreign language (EFL); see Section 2.1.3.1 The study of variation in SLA contexts, for further detail.

[2] This is essentially a phonetic variationist study; therefore, to obtain a trustworthy representation of the use of the variables, the use of phonetic transcription is required. The International Phonetic Alphabet (IPA), developed by the International Phonetic Association (also IPA), provides a clear guide to the standard representation of speech. From the last revision in 2005, two main charts are used in this study for referential purposes: the consonantic chart shown, in Figure i. and the vowels chart, shown in Figure ii; both are in Appendix A (for the complete IPA chart, refer to the International Phonetic Association, https://www.internationalphoneticassociation.org).

process, from first to fifth year. The interviewees are from both sexes and different socioeconomic groups. The aims of this study are:

I. To establish whether there are systematic patterns of variation in NNSs in the context of EFL – "the underlying systematicity of variable learner production" (Bayley, 2000: 288) – just as there are in native speech and non-native immersion contexts.
II. To identify potential similarities and differences between the patterns of variation exhibited by NSs and NNSs; that is, to test whether the variation in NSs of English is similar or different in NNSs, and to contrast the factors involved in the variation for these two groups of speakers.

To do so – and using traditional variationist methodology in the collection, manipulation, and analysis of the data – I examine the variation of four phonetic variables, each with a different number of variants.

1.2 The variables

The study of phonological features has traditionally uncovered great regularity in the patterns of variation as a consequence of social and linguistic regularities (Guy, 2007). In my study, the preliminary examination of the data (which is supported in *Section* 3.4.2.2 Coding*)* allowed me to select four different variables:

- Voiceless alveolar fricative (ʃ)
- Voiceless alveolar affricate (ʧ)
- Voiced dentoalveolar fricative (ð)
- Postvocalic (r)

Each of the variables is examined in relation to both internal and external factors (see *Section 3.4.3.2* below for an introduction to the factor groups included).

These particular variables were selected for a number of reasons: first, the first two variables (ʃ) and (ʧ) have not been widely studied in native speakers (NS) of English; their variation has been linked to other variables (e.g. in the case of (ʃ) as a variant for (s) in Scottish English, Stuart-Smith *et al.*, 2003) or as the alternation of (ʃ) and (tʃ) as a result of substrate influence (Labov, 2012); second, these two variables carry strong

social meaning in Chilean Spanish: the first includes among its uses an innovative variant characteristic of upper classes (Valencia, 1983-1984), and the second occurs as a stigmatised variant for the alveolar affricate /ʧ/, as such, characteristic of uneducated speakers from lower social classes (Cepeda, 2001; Rabanales, 2000; Valencia, 1983-1984). This research tests if the use of these variables in English reflects the social meaning carried by the use of these features in Chilean Spanish. The second two variables, (ð) and (r), have been extensively studied in sociolinguistics, and they have shown that the variation among NSs of English is constrained not only by social but also by internal linguistic factors (see e.g. Labov, 1962 for postvocalic /r/). As with the first two variables, (ð) and (r) have also shown to carry strong social meaning in L1 and L2 situations: for instance, the fortition of (ð), commonly referred to as *th-stopping,* has been described in dialects of English worldwide (Smith & Durham, 2011), many of them non-standard varieties (see e.g. Eckert, 2008); in the case of the latter, it has been commonly associated to ethnicity, and its status is associated to the ethnic group where it occurs.[3] Postvocalic (r), one of the most studied phonetic features in variationist research, also ranges from stigmatised to prestigious depending on the accent or dialect in which is present. A detail explanation that justifies the selection of (ð) and (r) is provided in *Sections 4.3 and 4.4*, respectively.

The decision to select these variables followed several considerations, amongst them: the amount of variation detected during the transcription process, which ensures that the feature is appropriate to be studied in terms of frequency and varying use (Labov, 1972), and second, the regularity of the patterns of variation associated with them, as demonstrated in previous sociolinguistic research. As mentioned briefly above, the four variables have shown different sociolinguistic valuations in the languages studied going from prestigious to stigmatised. Also, the literature has presented a considerable range of variants for the selected variables in both native and non-native speakers, which suggests that, in my study, this might also be evidenced. Thus, the variables are differentiated along a number of social and linguistic measures in both Spanish and English, allowing me to test for patterns of variability along these constraints.

[3] Eckert (2008) uses the concept "first order indexicality" introduced by Silvestein (2003), in which the use of the observed feature varies in status depending on the community that uses it; in other words, if the community that uses a particular variant is socially stigmatized, the feature will also be.

As this is a quantitative variationist study, I employed traditional variationist methodology to the actual design of the study (through the collection of the data by means of sociolinguistic interview – as defined by Labov, 1972) and to its analysis. Thus, over 14 hours of recorded interviews were obtained and transcribed, following the techniques described by Labov (1972), Milroy and Gordon (2003), and Tagliamonte (2006), amongst many others. Factor by factor and multivariate analyses are used.

1.3 Contents

The following chapters situate my research in theoretical terms; thus, Chapter 2 introduces the development and key concepts of variationist sociolinguistics, since it was pioneered by Labov (1963, 1966, 1972c, 1994, 2001) and followed by Cedergren (1973, 1987) and Valencia (1983-1984), amongst many others. This discussion is also complemented by general principles from SLA theory that could contribute to a possible explanation for the exhibited variation in L2 contexts, such as, e.g., the concepts of interference and interlanguage (Weinreich, 1953; Selinker, 1972; Tarone, 1983, 1985) and it reviews numerous studies in variation in SLA contexts, mainly to determine how the interaction of internal and external factors has contributed to the variation in L2 situations (Preston, 1989; Bayley, 1994; Major, 2004; Young, 1991; Preston and Bayley, 1996, 2005; Preston and Young, 2000). It also includes the study of variation in Chilean Spanish; this last section is intended to show the linguistic characteristics of the setting in which the data was obtained, including a phonetic characterisation of the first language of the speakers, as one of the aims of this study is to test for transfer.

Chapter 3 describes the methodology used during the collection of data as described in traditional variationist research (see, e.g. Labov, 1972; Tagliamonte, 2006), in what respects to the design of the research, the selection of the speakers, the collection of the data, and the transcription and coding processes. It also introduces and justifies the internal and external factors selected.

Chapter 4 consists of four sections, one per each variable chosen; each section reviews studies carried out using variationist methodology, followed by the analysis

of the results obtained from factor-by-factor and multivariate statistical analyses. Factor-by-factor analysis is conducted in all factor groups where possible.

Finally, Chapter 5 discusses the general outcomes of the research, comparing the results for the four variables examined in the previous chapter and providing the corresponding conclusions and limitations drawn from the present work. In addition, the final chapter indicates the possible implications of my study in SLA variationist research.

Thus, using traditional sociolinguistics methodology, this thesis aims to determine the source of variation in non-native speakers, and at the same time, compare whether this kind of variation is constrained by the same internal and external factors that condition native speakers'.

CHAPTER 2 - THEORETICAL FRAMEWORK

2.1 Introduction

As the primary focus of the present research is to uncover possible patterns of variation in Chilean speakers of EFL (henceforth NNSs of English), contrasting the variation shown by NSs of English, the logical approach is to set the theoretical basis on previous variationist research. This chapter will address the principles and relevant results described in variationist research in NS of English and Spanish, and from the interaction of both languages through the perspective of variation in second language acquisition research. I begin with an introduction to some of the key concepts that define variationist research, starting with a definition of the objectives of the study of language variation, through to the definition of variable, and an introduction to the preferred method of analysis.

2.1.1 Principles of variationist research

Based on the assumption that all natural languages vary and change (see e.g., Sapir, 1921, Weinreich *et al.,* 1968), the pioneering work of Labov in variation theory (see e.g., 1963, 1966, 1972) demonstrates that language variation can be explained by the effect of "external (social) and internal (systemic) factors" (Tagliamonte, 2006: 5). Its goal, as defined by Poplack (1993: 252), is "to discover patterns of usage, which pertain to the relative frequency of occurrence or co-occurrence of structures, rather than simply to their existence or grammaticality"; that is, to uncover patterns in the use of certain structures that cannot be explained solely through linguistic terms. According to Tagliamonte (2006: 5-6), this goal is founded on three facts about language: (1) the notion of orderly heterogeneity (Weinreich *et al.* 1968: 100), (2) the fact that language changes perpetually, and (3) that language conveys more than simply the meaning of its words. In other words, the inherent variation of language occurs in a systematic and continuous way, carrying more than purely linguistic information. To understand how language variation operates, she defines, as one of the key elements of any variationist study, the aim to *reach the vernacular* – (i.e., a "relatively homogenous, spontaneous speech, reserved for intimate or casual situations" (Poplack, 1993: 252). The vernacular is the style in which less attention is paid to speech, operating at the other end of the idealised norm (see also Labov,

1984; Milroy, 1992; Poplack, 1993; Tagliamonte, 2006), and which is obtained from the immersion of the researcher into a particular *speech community*. Thus, the element of analysis in variationist research is language itself, used by a selected group of individuals, which presents, as its smallest unit of study, the actual varying structure – the *linguistic variable*.

2.1.1.1 The linguistic variable

The linguistic variable occurs in the form of a choice made by the speakers at diverse language levels; this choice may take different forms to accomplish the same function (Poplack, 1993: 252); as defined by Tagliamonte (2006: 9), linguistic variables are "linguistic items which vary amongst themselves with the same referential meaning", or as is better put by Labov (1972), "two or more ways of saying the same thing".

The linguistic variable is, thus, defined as "the fundamental construct in variation analysis" (Tagliamonte, 2006: 70), and it is characterised by: (1) a high frequency of use,[4] (2) being a structural feature (it should be incorporated into a "system of functioning units"), and (3) its distributions should be highly stratified, in terms of, for instance, age levels or strata (Labov, 1972: 8). In other words, a linguistic variable must, firstly, be used with considerable enough frequency that its variation is a valid representative sample of the variation in terms of frequency of use – that is, that the number of occurrences is sufficient to cover a wide spectra of varying contexts; secondly, it must be a feature which is part of the functional structure of the language, that is, of one of the aspects, or levels, of language, such as the phonetic inventory; and thirdly, its distribution must be asymmetrical across the whole range of speakers, as an effect of the interaction between internal and external constraints, i.e., the use of a selected linguistic variable must be distributed unevenly in a sample which is made up of, for instance, people from different social classes, educational backgrounds, etc.

All the individual characteristics of the studied speakers play a role in the use of their language choices; however, the exhibited variation of a particular individual is often systematically patterned to other individuals of the same community – in terms of

[4] This particular characteristic may not be considered essential when defining a linguistic variable; instead, this can be considered a methodological concern, given that the high frequency of occurrence of a variable ensures enough data to conduct the analysis.

age, sex, ethnicity, educational level, etc. – in a more or less structured way (Poplack, 1993: 253). To these extra-linguistic factors, internal constraints also contribute to the election of linguistic forms. Thus, a variable is studied in conjunction with all the factors that may, or may not, have an effect on its occurrence. For instance, in the case of postvocalic (r), in his seminal study of New York department stores (1966), Labov studied the presence/absence of the variable *postvocalic (r)*[5] in three socially-stratified stores. He hypothesised that "if any two subgroups of New York City speakers are ranked on a scale of social stratification, then they will be ranked in the same order by their differential use of (r)" (Labov 1972: 44). This means that the presence or absence of postvocalic (r) in the speech of New York speakers would vary accordingly to their *social class*,[6] in addition to other (internal) constraints, such as phonological environment, style, etc.

2.1.1.2 The quantitative method

The inclusion of quantitative techniques in the study of language variation, since Labov's studies in several communities in the U.S., is based on the premise that variation is not random, and that the choices made by the speakers are systematically constrained by internal and external factors; thus, the main assumption in the quantitative approach is that this systematicity is accountable.

Quantitative analysis is a "combination of techniques employed in variation analysis" (Tagliamonte, 2006: 12), that comprises the "examination of individual instances of linguistic forms in the context of the grammar from which they come" (Tagliamonte, 2007: 190). In other words, by means of quantitative analysis it is possible to observe and quantify the behaviour of the varying instances under study in the contexts in which they occur.

According to Bayley (2002 [2006]), quantitative analysis allows the researcher to examine multiple combinations of factors (or cells); to do so, variationists usually employ statistical packages of the Varbrul kind (Cedergren and Sankoff, 1974; Guy, 1980, 1993; Young and Bayley, 1996). One of the most used tools in multivariate analysis is the programme Goldvarb, which is presented in different versions for

[5] He examined four occurrences of the use of postvocalic (r): casual fourth floor and emphatic *fourth floor* (Labov, 1972: 48; italics in original).

[6] Other factors examined are sex, age, race, and occupation, amongst others.

Windows OS and Macintosh (for a complete account on the use of Goldvarb, see e.g., Sankoff, 1988; Young and Bayley, 1996; Sankoff, Tagliamonte and Smith, 2005; Tagliamonte, 2006). While other statistical tools have shown they are inadequate, or can only relate an independent variable to a limited number of factors or dimensions (e.g., the use of t-test or analysis of variance; see e.g., Young and Bayley, 1996: 254), the "advantage of quantitative approach lies in its ability to model the simultaneous, multi-dimensional factors impacting on speaker choices, to identify even subtle grammatical tendencies and regularities in the data, and to assess their relative strength and significance" (Tagliamonte, 2006: 12). This means that, through quantitative analysis, variation is tested in relation to the constraining factors and how much of (and how) the effect of each factor contributes statistically to the variation it is also determined.

Quantitative analysis is often conducted in several steps: the first is the *distributional analysis*, in which the variables are examined across a series of cross-cutting factors, providing the "frequencies and percentages of the variant forms of the dependent variable in the data, either alone (an overall distribution of forms), or with the independent variables that condition or constrain it (a factor by factor analysis)" (Tagliamonte, 2006: 196). For instance, Tagliamonte (2006: 193-194) shows an example of distributional analysis, for the variable (t,d) deletion, which is displayed in Figure 2.1 below. This example is for the dependent variable variants "t", "0" and "?" for following phonetic environment vowels (V), consonants (C) and pause (Q) (bold in the original).

Group		t	0	?	Total	%
1 (new)						
V	N	456	44	20	520	45
	%	88	**8**	4		
C	N	227	240	31	498	43
	%	46	**48**	6		
Q	N	112	7	8	127	11
	%	88	**6**	6		
Total	N	795	291	59	1145	
	%	69	25	5		
TOTAL	N	795	291	59	1145	
	%	69	25	5		

Name of new cell file: ch9_eg2_28-3-05.Cel

Figure 2.1 Example of a distributional analysis, taken from Tagliamonte (2006)

In the example above, a total of 1,145 tokens (occurrences of the variable) are examined: 69% occurring with the variant [t] (N=795), 25% with the variant [0] (N=291), and 5% with variant [?] (N=59). It also shows that following vowels account for 45% of the instances (N=520), following consonants occur 43% (N=498), and following pauses 11% (N=127). Thus, for instance, Table 2.1 above shows that for variant [0], the most frequent following segment is a consonant (48% of the instances, N=240). This example can be applied to all factors considered in the design of the research, linguistic or social, as it shows "the fraction of times some particular outcome is found in the data" (Guy, 1993: 228).

Factor by factor analysis cannot test for the "genuine effect" of the factors examined or whether the variation is "due to chance" (Tagliamonte, 2007) and, therefore, *multivariate analysis* is conducted to "separate, quantify, and test the significance of the effects of environmental factors of a linguistic variable" (Guy, 1993: 237). When running a multiple regression analysis in e.g., Goldvarb, a step-wise series of procedures takes place to test for three lines of evidence: statistical significance, constraint ranking, and relative strength (Durham, 2006; Tagliamonte, 2006, 2007). In the first step (step-up), the group that makes the greatest contribution to the model is found; this process is replicated with all groups. Only those groups that are statistically significant to improving the model are added (Tagliamonte, 2006). This procedure is completed by a reverse analysis (step-down) in which all groups are

initially included and then those “whose loss least significantly reduces the likelihood” are discarded (Tagliamonte, 2006: 143; Durham, 2006).

The results drawn from the step-up/step-down analysis are then displayed in a table, as in Table 2.2, elaborated for [t,d] deletion by Tagliamonte (2006: 247-248). All statistical values (the weight of each factor) range from 0 to 1, in which “weights below 0.50 […] ‘inhibit’ the operation of the rule and those above that weight ‘promote’ it’” (Preston, 1996: 10), that is, from the example below, the factor weight for “following phonological segment: obstruent = .83” would *favour* the application of [t,d] deletion (rule), and “following phonological segment: vowel = .30” would *disfavour* the application of the rule.

Table 2.1 - Example: "Multivariate analyses of the contribution of internal and external factors selected as significant to the probability of [t, d] deletion; factor groups not selected as significant in square brackets", adapted from Tagliamonte (2006, pp. 247-248).

		Contemporary British English		
	Corrected mean			.17
	Log likelihood			-430.787
	Total N			1232
		Factor weight	%	N
Following phonological segment	Obstruent	.83	52	357
	Glide	.70	37	111
	/r/	.60	25	32
	/l/	.50	23	26
	Vowel	.30	8	570
	Pause	.20	5	136
	Range	63		
Preceding phonological segment	Sibilant	.69	40	367
	Nasal	.45	17	439
	Liquid	.34	21	130
	Stop	.43	16	169
	Non-sibilant fricative	.29	12	127
	Range	40		
Morpho-logical class	Monomorpheme, e.g., mist	[.53]	26	716
	Irregular past, e.g., kept	[.50]	21	128
	Regular past, e.g., missed	[.45]	19	388
Speakers', sex	Male	.59	30	484
	Female	.44	22	634
	Range	15		
Speakers' age	14-24	[.45]	24	375
	35-44	[.54]	24	272
	55-64	[.55]	27	262
	75-84	[.48]	27	209

In the table, several other values are incorporated:

- The *corrected mean* or *input* corresponds to the "overall indication of the strength of the rule"; in the example above, .17 would indicate that the rule "deletion" is not very likely to occur. However, when it does occur, its variation is constrained by the factors, groups following and preceding phonological environment, and speaker sex (the significant factor groups).
- The *range*, obtained by subtracting the lowest from the highest factor weight value, indicates the effect of the factor in the application of the rule; the higher the value, the greater the contribution to the variation; thus, in the example above, following phonological environment is the strongest contributor to the application of the rule and speaker sex the least.

The use of multivariate analysis is, however, conditioned to a relatively high number of tokens and to the non-existence of empty cells (e.g., the non-occurrence of one of the variant in one phonological environment). When multivariate analysis cannot be completed, factor-by-factor analysis can suffice to determine the effect of the internal and external factors on the variation of a particular variable. This means that multivariate analysis requires that all cells, i.e., context or factors, present non-categorical variation, as otherwise they would be invariant.

The results drawn from factor-by-factor and multivariate analyses do not only provide the researcher with solid evidence of the effect of a particular factor – or interacting group of factors – on a specific variable; they also serve the purpose of allowing the researcher to compare, in quantitative terms, the patterns of variation of different data sets, from different linguistic communities (Tagliamonte, 2007).

Having introduced the most relevant concepts that have given origin and support to the study of language variation, I now turn to the most relevant studies that provide the theoretical basis for this particular research.

2.1.2 Variationist research in English

Since the 1960s, most of the fundamental literature in variationist research has dealt with L1 situations, in this case, with native speakers of English speaking English. The authors that have set the basis for many both theoretical and empirical works are Labov (*Martha's Vineyard*, 1963 and *Social Stratification of English in New York City*, 1966), Trudgill (*Social Differentiation of English in Norwich*, 1974), Cheshire (*Variation in an English Dialect: A Sociolinguistic Study,* 1982), Wolfram (*A Sociolinguistic Description of Detroit Negro Speech*, 1969), Milroy and Milroy (*Belfast: Change and Variation in an Urban Vernacular*, 1978), amongst many others. The aim of the present section is twofold: first, to provide an overview of the most relevant findings in variationist research that are directly related to the variables examined, and second, to provide my own study with the elements of contrast for the linguistic variation of NNS against NS of English.

The study of language variation, as described by Labov in the series *Principles of Linguistic Change* (1994, 2001, 2010), has shown that the conditioning factors in the varying use of language can be classified in three groups: internal, external, and cognitive and cultural; here I focus only on the first two groups for two of the variables studied in my work: postvocalic (r) and voiced dental fricative, as an introduction to the study of variation in English.

Social class has been, perhaps, the most studied factor in many sociolinguists' investigations. This is particularly the case with postvocalic (r), where social class tested to be the most relevant conditioning factor in the use of this variable for both New Yorkers and black speakers from Detroit (Labov, 1966 and Wolfram, 1969, respectively). A particularly interesting phenomenon that arises in Labov's study is hypercorrection,[7] in which middle class speakers perform over upper class speakers in more formal situations (reading list and minimal pairs). This factor can be seen, in the opposite pattern, in Britain, where the absence of the variable is not connected to lower social classes, but to upper (see e.g., Trudgill, 1974; Foulkes and Docherty, 2000). I return to these issues in Section 4.4.2.1.

[7] This phenomenon is examined in detail in Chapters 4 and 5.

In relation to another external factor, *sex*, Wolfram's study revealed another very relevant fact for variationist sociolinguistics: that female speakers use more standard features than male speakers; this finding will later become one of the principles of the gender paradox, in which, in stable linguistic situations, women generally use more standard features than men, and in language change situations, they (a) (in change from above)[8] are the leaders in the change of linguistic features, at the same time, and (b) (in change from below) use the new variant more than men (see Labov, 1990, 2001).

In both Wolfram and Labov's studies, *age* tested to correlate with language variation, especially in Wolfram's study; here, the adults used more stigmatised variants than teenagers, but this factor intersected with social class, confirming the predominance of class over other factors.

Style tested to be another highly relevant factor in the use of standard versus non-standard features for both Labov´s and Wolfram's works, in which more formal styles tend to favour the use of standard or prestigious features and more informal or casual situations favour the use of non-standard or stigmatised features. These two research investigations have since been replicated by numerous studies, incorporating new factors and methodological aspects; among these are Levine and Crokett, 1966; Anshen, 1969; Myhill, 1988; Feagin, 1990; Bailey, *et al.*, 1991; Bernstein, 1993; Bayley and Thomas, 1998; Labov, 2001; Schönweitz, 2001; Starks and Bayard, 2002.

Voiced dental fricative (ð)[9] has also been widely examined since Hubbell's 1950 study in New York.[10] His findings, especially the use of stop variants correlated to class, were later replicated by Labov (1966a) in New York, Shuy, Wolfram and Riley (1968a) and Wolfram (1969) in Detroit, Anshen (1969) in North Carolina, Milroy and Milroy (1978) in Belfast, Winford (1978) in Trinidad English, Dubois and Horvath (1998) in Cajun English, Kerswill and Williams (2000) in Milton Keynes, Reading and Hull, Stuart-Smith (2007) in Glasgow, Eckert (2008) in New Jersey, and Gibson and

[8] The concepts of *change from above* and *change from below* refer to the level of consciousness that speakers have about the linguistic change of a given linguistic variable; in *change from above*, the speakers are aware of an ongoing linguistic change and in *change from below*, they are not (see e.g. Labov, 1966, 1972)

[9] The third variable under study here.

[10] Labov (2001) reviews descriptive studies conducted in New York prior to Hubbell"s, but these are based more on observations, without using variationist methodology (Babbitt, 1896; Frank, 1948).

Bell (2010) in New Zealand, amongst others. One of the main constraining factors on the use of this variable was found to be gender, with women favouring the standard variant over male speakers; as this variable is considered to be a stable sociolinguistic variable, this gender-oriented linguistic behaviour has been justified under the gender paradox, mentioned above.[11] This research is not a study of language change or of a language in a stable situation, but rather a consideration of the NNS of English and of the potential effect of their L1 on their linguistic performance; however, studies in these areas provided me with a theoretical framework, not only in relation to the results obtained in the different communities of NS speakers of English, but, most importantly, on the methodology used to uncover linguistic patterns – from the design of the studies, to the collection and analysis of the data, and the subsequent display of the results.

Just as variationist research has been conducted in L1, during the last two decades it has also been the object of interest in L2 studies, as, in the words of Preston and Bayley (1996: xiii), "variability is among the most characteristic features of learner speech" and "few studies of interlanguage variation have made use of the methods of multivariate analysis of linguistic variation". My study attempts to fill, in part, this gap in the literature, especially concerning the transfer of sociolinguistic patterns from the L1 to the acquiring language. Thus, the following section introduces those studies that employ variationist methodology in the study of the speech of NNS.

2.1.3. Second language variation

As shown in the previous section, the variation of L1 speech has been widely studied for English, but in the context of my research – variation in L2 acquisition – and despite the efforts made in the last decade, the number of studies conducted using the quantitative methodology is not considerable. This section reviews some studies that have examined L2 variation, i.e., how non-native speakers' variation can be correlated to the factors commonly studied in L1 variation situations. Thus, here I provide an insight into some general concepts and theories that will help us understand how variation has been justified in EFL situations.

[11] See pp. 32-33 this same section.

According to Ellis (1997: 4-5), the two main objectives of SLA are the description and explanation of the acquisition of an L2, and, within the latter, the identification of the external (input) and of the internal factors (cognitive mechanisms) involved in acquisition.[12] Some of the key issues in SLA research have been: the definition of L2, the difference between *learning* and *acquisition*, the form of acquisition of the different levels of the language, the age of the acquirer, the transfer of structures from the L1 to the L2, the role of the L1, the immediate linguistic context, individual differences, and the existence of universals, amongst others.[13] As the focus of the present research is on the variationist perspective of SLA, I will not discuss theoretical issues concerning the acquisition of second or foreign languages.[14]

If we go back to the definition of the objectives of SLA research provided by Ellis (1997) above, there are at least two aspects left aside: the context of acquisition[15] – and how this may impact on the language acquisition process – and the potential systematicity of the variation exhibited by non-native speakers; that is, if native speakers vary (systematically) in relation to a series of factors (and, within these factors, the effect of their social context seems relevant), do non-native speakers also vary?

One of the key concepts that helps to answer these questions and incorporate the social dimension into the study of language acquisition is that of *communicative competence* provided by Campbell and Wales (1970). They suggest that, to the linguistic competence usually associated to the optimal use of grammatical rules, contextual or sociolinguistic competence should be included. This sociolinguistic

---

[12] Other theories, such as Universal Grammar (UG) (Chomsky, 1965, 1993; Pinker, 1994), suggest that the acquisition of a language is a biologically bestowed faculty, which will explain why subjects "come to know properties of grammar [i.e., of the language system] that go far beyond the input" (White, 2003: 20). The discussion on whether UG, or any other theory, explains the processes operating behind L2 acquisition are out with the scope of this research.

[13] Most of these topics are discussed extensively in, e.g., Krashen (1981), (Klein, 1986), Doughty and Long (2008) and Gass (2013), to name a few.

[14] It is out with the scope of this research to discuss the distinction between learning and acquisition, just as it is to try and explain the variation in function of a specific theory of SLA.

[15] By "context of acquisition", I refer to the social setting in which language is spoken, which is different from the context of acquisition studied in traditional SLA research. In the first, context may or may not have an impact on the acquisition of sociolinguistic competence (e.g. in the case of immigration settings), whereas the latter has typically focused on the description of the L1 system (language) as context and how this can have an effect on the acquiring system (e.g. how the structures of Spanish may have an effect on the acquiring structures of English).

competence implies knowledge of the rules of language use (Canale and Swain, 1980: 4), i.e., the learner must not only know the language as a system, but also must know how language works in its context.

Thus, SLA research incorporates the study of the variation in the interlanguage[16] of the learners of L2, focusing initially "on socially marginalised speakers, often working-class immigrants" (Bayley, 2007: 133), such as the case of Schumann's 1978 study on *Alberto* (a Puerto Rican immigrant to the U.S.) and Huebner's 1983 study on *Ge* (a Hmong immigrant to Hawaii). As the systematicity found in NSs was also tested in NNS, variationist studies were extended to other communities, such as in Dickerson (1975) on Japanese students of English in the U.S.

The examination of second language variation has determined the existence of two types of non-native variation (see e.g. Regan, 1996; Rehner, Mougeon and Nadasdi 2003; Dewaele, 2004a; Bayley and Regan, 2004, Mougeon, Rehner and Nadasdi, 2004; Durham, 2006; Mougeon, Nadasdi and Rehner, 2010; and Howard, Mougeon and Dewaele, 2013):[17]

- *Type I* variation, which is defined as "the learner's variable use of two or more L2 forms to express the same functional value, where one or all forms are non-native" (Howard *et al.*, 2013: 1). In other words, the learner produces a variable range of forms that may not correspond to the target. This variation is characteristic of earlier stages of the language acquisition process and it decreases as the learner is exposed to the language (*ibid*.). Other authors have referred to this kind of variable interlanguage production as variation along the vertical continuum (Corder,1981; Young, 1988).[18] A few examples of this kind of variation are the studies conducted by Adamson (1989), Bayley (1996), Young (1996), Beebe (1977), Wolfram (1985), Noriega (2004), Ellis (1987). From these studies, one of the most relevant findings suggests that

[16] The concept of interlanguage is discussed in Section 2.1.3.1.2.

[17] Durham (2006: 103-106) provides the terms interference-based variation (IBV) for Type I variation, and target-based variation (TBV) for Type II.

[18] The vertical/horizontal distinction refers to the assumption made by Corder (1981: 91) that, while in the initial stages of the acquisitional process the learner varies vertically along the scale of complexity, ("within the range of options he has available", e.g. by means of code switching, backsliding, foreign, or baby talk), more advanced learners vary across it (see also Regan, 1995: 178).

the phonetic environment is a relevant conditioning factor, just as in NSs speech (Wolfram, 1985; Noriega, 2004).

- Type II variation, which "is concerned with the learner's ability to use native-speaker sociolinguistic and socio-pragmatic variants" (Howard *et al.*, 2013: 1). This kind of variation supposes a higher level of competence in the L2 than Type I, probably as it "only emerges following stability in the learner's grammatical development" (Howard *et al.*, 2013: 2; see also, Adamson and Regan, 1991). Furthermore, Bayley and Regan (2004: 325) suggest that "in order to become fully proficient in the target language, second language learners also need to acquire native-speaker patterns of variation". Numerous studies show that, despite the fact that the use of native-speaker-like sociolinguistic variation increases over exposure (e.g. in naturalistic contexts), standard or formal variants are still preferred over informal or vernacular variants (see Howard *et al.* 2013 for a complete account of these studies). The earliest works that examine this type of variation are Anisman (1975) and Thompson (1976), who studied how ethnicity and social class had an effect on the choices made by Puerto Rican and Chicano speakers of English, respectively. Other studies that fit in this category are Adamson and Regan (1991), Bayley and Preston (1995), Rehner, Mougeon and Nadasdi (2003), Mougeon Reehner and Nadasdi (2004), Wolfram, Carter, and Moriello (2004), and Hansen (2001, 2004), amongst others.

While Type II variation usually occurs in language contact scenarios (e.g. *naturalistic environments*: Blondeau *et al.*, 2002; Adamson & Regan, 1991; Mougeon and Rehner, 2001; Bayley, 1996; Major, 2004; *immersion settings*: Mougeon, Rehner and Nadasdi, 2004; Nadasdi, Mougeon and Rehner, 2005; Mougeon, Nadasdy and Rehner, 2010; *naturalistic vs. classroom*: Howard, 2010; Regan, 1996; Barron, 2003; Mougeon and Rehner, 2001; Rehner, 2010, 2011), Type I variation is characteristic of classroom environments (see e.g. Dewaele 1992, 2004a, 2004b).[19] This study should be considered with the latter of these. Although most previous studies have dealt with advanced L2 learners, my research focuses on three stages of the acquisition process: beginner, intermediate and advanced learners of English, so providing second language variation literature with one, relatively neglected, aspect of non-native

[19] See Section 2.1.3.2 for an account on studies in naturalistic, immersion and classroom environments.

variation. As in Durham (2006: 103), establishing a clear difference between Type I and Type II non-native variation will help me to formulate secondary hypothesis for each variable, depending on the status of these in the two languages.[20]

In terms of theory that connects variationist research and SLA, Preston (1989) provides a full account of the matters that connect both linguistic areas. For instance, he examines the variation of NS (from Labov, 1966 onwards) in terms of ascribed and acquired characteristics, and then extends the findings to studies of NNSs acquiring a L2 in different contexts (immigration, instructional, and language contact situations – pidgins and creoles). Similarly, Bayley and Preston (1996) offer a collection of contributions in an attempt to "correct the relative neglect in SLA" studies from the variationist perspective, going from a description of quantitative analysis concepts and methodology to the inclusion of several studies conducted in immigrant and instructional settings.

As mentioned in the introduction of this section (pp. 35-36), the discussion of theoretical aspects of SLA is not the concern of my research; however, it is important to review how variation has been seen historically from the SLA perspective. Thus, in the following section I summarise only those concepts that are relevant to understanding the change from the traditional SLA concept of acquisitional *mistake* to acquisitional *variation*.[21]

2.1.3.1 The study of variation in SLA contexts

From the earlier studies on SLA, two key issues have been at the centre of the debate: whether L2 structures in any of the different levels of language intermix with L1 structures (for language transfer and interference; see Weinreich, 1953) or if they keep separate, forming new interlanguage systems (for the Theory of Interlanguage, see Selinker, 1972 or refer to Section 2.1.3.1.2 below); these two SLA theories are

[20] Durham (2006) proposes three possible cases of variability: Case A, where the variation is acquired (TBV); Case B, where the outcome is a focused form (or common problems of acquisition); and Case C, in which variation is only influenced by the native language. See Section 2.1.3.2.2.1 for a complete explanation of this model, and how the difference between the two types of variation can be applied to this group of non-native speakers.

[21] In most SLA theories prior to the study of variationist SLA, the alternative forms used by the speakers were seen as mistakes, occurring with no identifiable pattern (see e.g. *Error Analysis Theory*, developed from Corder's 1967 *The significance of learners' errors*).

reviewed in the following sections. What has been agreed upon, however, is that the process of acquisition is, by itself, a social process (see e.g., Fillmore, 1991; Tomasello, 1992; Ellis, 1997), resulting from social interaction, whether this be in a second language or in a foreign language context, in instructed education or by means of immersion, in monolingual, bilingual, or multilingual societies.

2.1.3.1.1 Language transfer and interference, and the role of perception

In his seminal work *Languages in Contact,*[22] Weinreich (1953: 1) defines interference as "those instances of deviation from the norms of either language (L1 or L2) which occur in the speech of bilinguals as a result of their familiarity with more than one language, i.e., as a result of language contact". Interference, then, in the words of Weinreich, implies the "rearrangement of patterns that result from the introduction of foreign elements into the more highly structured domains of language, such as the bulk of the phonemic system, a large part of the morphology and syntax, and some areas of the vocabulary". This means that interference is the deviant use from the norm, as exhibited by bilinguals as a result of the interaction of the phonemic, morphological or syntactic systems of both languages, mainly as the L2 contributes with new elements to the L1. He distinguishes between negative transfer (interference) and positive transfer, which is the facilitating influence of the similarities of one language onto the other.

Although Weinreich defines interference (positive and negative) for bilingual speakers, Odlin (1989: 27) uses this term in the context of L2 speakers, defining it later as "the influence resulting from the similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired"; here, the variation seems to result from the structural differences (or similarities) of both L1 and L2, influencing each other. A third definition is provided by Flege (1995: 233), who suggests that [phonetic] interference is the re-adaptation of the L1 phonetic system in relation to the sounds "encountered" in the new phonetic system, i.e., the L2, results in what can be recognised as a "foreign accent". In all these definitions, the common ground is the fact that when two

[22] The work of Weinreich is based on bilingual and multilingual situations; the difference between bilingual and L2 situations is based on the first having two mother tongues (or more in the case of multilingual communities) and the second, the acquisition of a second language occurs either by immersion or in an instructional setting (EFL, for instance).

languages interact, either in bilingual or L2 situations, there is, to some extent, a positive or negative transfer of structures from the dominant language (or L1 in the case of L2 situations) into the second language.[23]

Specifically, in terms of phonic interference Weinreich (1953) suggests four types of interference:

- Under-differentiation of phonemes: "two sounds of the secondary system (L2) whose counterparts are not distinguished in the primary system (L1) are confused" (18); that is, the learner does not note the difference between one sound of their L1 and two sounds of the acquiring language.
- Over-differentiation of phonemes: "involves the imposition of phonemic distinctions from the primary system on the sounds of the secondary system, where they are not required" (*ibid*.). Here, the sounds from the L1 appear in the L2, even when they are present in the phonetic inventory of the latter.
- Reinterpretation of distinctions: "the bilingual distinguishes phonemes of the secondary system by features which in that system are merely concomitant or redundant, but which are relevant in his primary system" (*ibid*), that is, the individual identifies a subordinate sound in the L2 by associating it to a primary sound in their L1.
- Actual phone substitution: "applies to phonemes that are identically defined in two languages but whose normal pronunciation differs" (*ibid*); this means that the use of the phonemes differs in L1 from L2, despite the fact that they are described as the same in both phonetic inventories. Weinreich suggests that this kind of substitution "is a type of interference in which the identity of other phonemes is not disturbed" (19, also note 9), and it refers to "features which are redundant but which are apt to become relevant if the phonemic system changes" (*ibid*); in other words, these are phonemes whose status is changeable. Weinreich also refers to this type of phoneme substitution as "the non-customary pronunciation of an identifiable phoneme."

[23] There are different types of bilingualism: compound, coordinate, and subordinate. In the first, the individual learns both (or more languages) at the same time, in the same environment; in the second, the speakers acquire the two languages in two different settings, (it could be, in the case of immigrants, at home and in school); and in the third case, there is a dominant language. In relation to the status of the languages, Weinreich (1953: 74-75) suggests that, from a psycholinguistic criterion, there is a dominant language, whose status is subject to change based on proficiency, order of learning, and attitudinal criteria.

The observation of the occurrence of the above mentioned phenomena has also been studied in L2 speakers, for instance in relation to the role of perception (see e.g., Flege, 1995 for native speakers of Italian learning English in an immersion setting). Studies suggest that when this production diverts from the norm, it is considered an *error* (see e.g., Flege 1995 for a complete account on this); for instance, based on the idea that native-like language production is trainable (without considering the occurrence of alternative forms), Flege (1995: 237) and his colleagues have developed a speech learning model (SLM), which allows the learner "to articulate or imitate proficiently foreign sounds". In this model, perception is a key function in the acquisition of foreign forms. The difficulty to produce specific sounds might be due to the speaker's failure to distinguish the differences "between pairs of sounds in the L2, or between L2 and L1 sounds. This difficulty occurs either because phonetically distinct sounds in the L2 are 'assimilated' to a single category, because the L1 phonology filters out features (or properties) of L2 sounds that are important phonetically but not phonologically, or both" (*ibid*.). This misperception of features might not be permanent and the speakers' ability to produce L2 sounds is likely to improve due to exposure by means of "naturalistic learning" (ibid.).[24] In other words, the production of native-like forms is achievable with the passing of time: the more the learner is exposed to English, the closer they will be to native-like production.[25] These sorts of findings are potentially very important for my research, as they suggest that, firstly, the production of a sound is a mirror of its perception, and, therefore, the realisation of a particular feature is the result of how it is heard. Also, it attributes the similarities or differences between the first language (L1) and the second (L2) to perception and not only to production or the acoustic characteristics of a phonetic system, such as the "alleged" missing phonemes in Spanish (for this particular claim, refer to the variable voiced dental fricative, in Chapter 4.3).

One aspect of Flege's SLM worth noting is that, amongst the many factors considered in the native-like production of L2 sounds (at least in the cas of immigrant settings) are: age of acquisition, the quality of the input received by the individual, length of

[24] In this context, 'naturalistic learning' refers to the use of the environmental setting to learn/acquire features from the L1 and not in the pedagogical sense defined by Gardner (1983).

[25] In the case of the speakers of my sample, they receive the input not only from their instructional setting, but most probably from the media, by means of music, TV, cinema, internet, etc.

residence,[26] age of arrival, motivation and economic and social factors (see e.g., Piske, MacKay and Flege, 2001 for a full account of studies on these topics), linking the use of L2 forms not only to purely linguistic (acoustic) properties of the features, but also to social factors.

The study of interference has provided SLA with one theory to explain the variation found in bilinguals and non-native speakers acquiring a second language, as it considers the alternative realisations of elements of the language as the result of the first (or dominant) language facilitating (positive transfer) or interfering with the second language (negative transfer or interference). This theory will be referred to again in the Chapter 5. The following section deals with another theory that attempts to explain variation from the SLA perspective.

2.1.3.1.2 Interlanguage (IL) theory

Whereas in interference theory the two (or more) language systems stay separate, interlanguage refers to a new emerging system occurring in speakers learning a second language. Selinker (1972: 214) defines it as a "separate linguistic system based on the observable output which results from a learner's attempted production of a TL (target language) form". He presents the notion of TL as "the second language the learner is attempting to learn" particularly restricted to "only one norm of one dialect within the interlingual focus of attention of the learner" (213); this means that during the process of acquisition of a second language, the learner develops a new, intermediate system, based on the norms of the target language, but restricted to one variety of said language. A key outcome of interlanguage is the concept of fossilisation, which Selinker (1972: 215) describes as the maintenance of certain "linguistic items, rules, and subsystems [...] in the IL relative to a particular TL", i.e., the mistakes made by the speaker do not belong to the NL[27] (natural language or L1) system but to the IL. These structures, argues Selinker (1972: 215), "tend to remain as potential performance, re-emerging in the productive performance of an IL even when seemingly eradicated." This means that every deviant use of a particular feature

[26]Flege's 1998 study on Italian and Korean speakers of English as a foreign language considered individuals resident in Canada and the U.S., respectively.

[27] Here, Selinker does not distinguish between the terms 'natural language (NL)' and 'first language (L1)' to refer to the mother tongue of the speaker.

(in the case of this research, a sound) is a fossilised form that originates in the intermediate system (the IL) as an effect of the L1.

In the formation of this separate linguistic system, i.e., the IL, Selinker (1972: 215-220) describes several processes that may take place:

- Language transfer: when "the fossilizable items, rules, and subsystems which occur in IL performance are a result of the NL"; that is, the items, rules and subsystems are transferred from the L1 to the IL.
- Transfer-of-training: when "the fossilizable items, rules, and subsystems are a result of identifiable items in training procedures"; e.g. the non-distinction on the use of PN he/she in Serbo-Croatian learners of English, as "they feel they do not need to make this distinction in order to communicate". This means that, during the acquisitional process, the individuals choose an (alternative) use of a particular feature as an influence from their L1.
- Strategies of second-language learning: when "the fossilizable items, rules, and subsystems are a result of an identifiable approach by the learner to the material to be learned", e.g. the tendency, on the part of learners, to reduce the TL to a simpler system (all-verb-are-transitive sort of simplification). This means that the learner utilises an item rule or subsystem as a way to facilitate the learning of them.
- Strategies of second-language communication: when "the fossilizable items, rules, and subsystems are a result of an identifiable approach by the learner to communication with native speakers of the TL", e.g. "the tendency [...] to avoid grammatical formatives such as articles, plural forms, and past tense forms". Here, the use of deviant forms obeys an attempt by the learner to communicate efficiently.
- Over-generalisation of TL linguistic material: when "the fossilizable items, rules, and subsystems are a result of a clear overgeneralization of the TL rules and semantic features", e.g. "extension of –ed (past tense morpheme) to contexts other than affirmative." In this case, the use of deviant forms occurs as a consequence of an overgeneralisation of rules and other features from the L1.

All the processes above involve the notions of deviant forms, used primarily by the learner to perform adequately in the L2; therefore, the learner uses the fossilisable items, rules and subsystems stored in their IL. Selinker (1972: 217) goes even further, arguing that "entire IL competences (can) be fossilized in individual learners performing in their own interlingual situation, but also in whole groups of individuals, resulting in the emergence of a new dialect, where fossilized IL competences may be the normal situation". He asserts that it is due to fossilisation than we can recognise accents such as, e.g., Indian English. This would mean that, if the features found in my data characterise the speech of Chilean speakers from other speakers of English as a foreign language, it can also be attributed to fossilisation and not only to transfer; however, the design of the research does not allow me to test whether or not the variation is the result of IL processes (nor to test for transfer), considering that the aim of it is to test that the source of variation is in the interaction of social and linguistic factors. I will return to this issue in the discussion in Chapter 5.

Given that the focus my research is in variationist linguistics, the inclusion of discussion of these two theories in this chapter stems from two main facts: firstly, they provide the basis to study the variation in L2 speakers, as they acknowledge the occurrence of deviant forms in the process of acquisition of the L2; as many of the varying features in L2 literature (either explained by transfer or fossilisation) are patterned across languages (e.g., in morphology, third person singular agreement in Arabic (Al-Badawi, 2012), Chinese (Goad, White, and Steele, 2003; Hawkins and Casillas, 2008), Turkish (White, 2003) and Spanish (Hawkins and Casillas, 2008), amongst others), it can be argued that the variation of non-native speech is also structured, just as variationist research has tested for native speech. Secondly, both interference and IL theories offer different explanations to the variation, specifically of what occurs during the acquisition process; thus, seen from an acquisitional point of view, the variation resulting from the interaction of social and linguistic factors can find, at least within IL, a contributing explanation in the process of acquisition of a second language. The effect of the L1 on the L2 as a potential factor will be considered in Chapter 5.

As my research is centred on Chilean speakers of English as a foreign language, the examination of variation in L2 is essential, and, therefore, some of the key findings for NNS variation are presented next, in both immersion and EFL settings.

2.1.3.2 From theory to practice

As shown in Sections 2.1.3.1.1 and 2.1.3.1.2, variation in NNS has largely been explained through L1 interference (Weinreich, 1953; Krashen, 1988; Odlin, 1989; Flege, 1995) or IL Theory (Selinker, 1972). Although these undoubtedly play a part in the use of non-target forms, sociolinguistic influences may also be important, as suggested by Schmidt (1987) for Egyptian Arabic speakers of EFL, particularly the effect of style and socioeconomic class, and Beebe (2006) for Thai speakers of ESL, for the factors of style and socioeconomic class. The following section presents the main constraining factors in the variation of non-native speakers in immigrant and immersion settings, followed by the findings for non-native speakers in instructional settings (EFL).

2.1.3.2.1 Variation in non-native speakers in immigrant and immersion settings

The study of variation of speakers who acquire a L2 in the place where the L2 is spoken as L1 (e.g., Poplack, 1978 and Wolfram, 1971, for dialectal acquisition in bilingual Puerto Rican speakers of English in the U.S.), has shown that the transfer of, for instance, phonemes from L1 to L2 is constrained by external factors,[28] such as time of exposure to English (Flege, Munro and MacKay, 1996; specifically on age of arrival, therefore, time of exposure) or network (Bayley, 1996), or by internal constraints such as the preceding (Carlisle, 1997) or following segment (Benson, 1988). The number of studies that report on the transfer of sociolinguistic patterns from L1 to L2 (as in my study) are rare; for instance, Bayley (1996) studied [t,d] deletion in Chinese L1 learners of English,[29] finding differences in the acquisition of native and non-native speakers of English, constrained by internal (grammatical category, preceding and following segment, style) and external factors (social network, proficiency); one very important conclusion from this study suggests that those non-native speakers (NNS) that interacted more with native speakers (NS)

[28] The distinction between external and internal factors is examined in detail in Section 3.4.3.2 Factor Groups; however, at this point, I will use 'external factors' to refer to the group of external constraints to the language – such as sex, education, socioeconomic class, etc. – and 'internal factors' to refer to those constraints that are motivated from inside the language. These are usually referred to as *social* and *linguistic factors*, respectively.

[29] In an immersion setting, over periods ranging from two months up to five years of residence in the U. S..

acquired similar patterns of variation to NS than those who used English in instructional contexts only, which means that the effect of social interaction is a relevant factor to be considered in the study of NNS variation.

2.1.3.2.1.1 Main findings for the variation of NNS in immersion settings

The decision to include the results of numerous studies on the variation of non-native speakers of English in immigrant settings follows the idea that the examination of this sort of pattern may help us to determine the nature of variation in NNS (source of variation and constraining factors) and how it is different, or similar, in immigrant and foreign language settings. Some of the factors examined in the study of SLA variation are described next:

1. *Preceding and following segment*: Bayley (1996) reports on the widely studied variable t/d deletion, in twenty native speakers of Mandarin living in the U.S. He found that, just as is found in native speakers of English, -t/d was less "likely to be deleted following a liquid than an obstruent or nasal" (103-104), but unlike native speakers, there is no difference in the effect of preceding obstruents and nasals; for following segments, the deletion hierarchy is the same for native and non-native speakers: consonant > glide > vowel. He reports some differences in the results drawn by Labov's 1989 study, as -t/d is "slightly less likely to be deleted before pauses than before vowels" (104).
2. *Gender*: Major (2004: 173) examines the acquisition of gender and stylistic differences in NS of English against NS of Japanese and Spanish, acquiring English as a L2; he studies "four widespread casual speech processes occurring in American English: (1) palatalization (e.g., *got you* → *go[č]you*), (2) deletion of *v* in *of*, (3) -*ing* versus -*in*', and (4) *n* assimilation in *can* (e.g., *he ca[ŋ]go*)"; one of the most relevant conclusions drawn from Major's study indicates that gender differences are acquired before stylistic differences, therefore the inclusion of gender as a factor – as in many variationist studies of NS – has tested to be a relevant factor in NNS and calls for its inclusion in my own analysis.
3. *English proficiency:* Bayley (1996) considers as a constraining factor in the deletion of –t/d (see point 1 above, on preceding and following segment) the

proficiency of the speakers in their L2.[30] His results show that the highest rates of –t/d deletion is found among lower proficiency speakers, as he predicted: the higher the proficiency of the speakers, the more they mark the verbs for tense and the more they mark final clusters. If extended to my research, his results would mean that the more advanced students would show higher rates of target-like features; however, it can only be compared in general terms, as this feature has shown different patterns of variation for native and non-native speakers (see point 1 above) and second, it is not a feature considered here.

4. *Differences in the phonetic inventory of the interacting languages*. Flege, Munro and MacKay (1996) also consider the possibility that the phonetic differentiation of consonants is the result of the absence of a given sound in the L1 inventory, i.e., that those phonemes that are not in the phonetic inventory of the L1 are perceived, and therefore produced more effectively than those which are common for both languages. The results were not conclusive with this group of speakers, but a study on perception conducted by Morosan and Jamieson (1989) with native speakers of French learning English suggests that it was difficult for the group to make differences between the phonemes /d/ and /ð/ and /t/ and /θ/, in which the second of each pair is not present in French. These results are interesting for my research as there is a difference between the English and Spanish phonetic inventories in relation to the four variables studied.[31] This factor will be reviewed in the discussion section, but it will not be included in the analysis section given that the above mentioned studies are of a different nature, focused on perception, and not in the actual use of the variables.
5. *Age of acquisition*. Flege, Munro and MacKay (1996) examined the production of word-initial English consonants by a group of Native Italian immigrants in Canada, and contrasted the results with native speakers of English. Among the hypothesis they wanted to test was that the age of acquisition was a major

[30] He considers two groups: high proficiency and low proficiency speakers. The first is made up of speakers with TOEFL scores of 550 or higher, the second of speakers with TOEFL scores of 510 or lower, and pre-TOEFL.

[31] The four variables studied here have different values in English and Spanish; for instance, one of them, postvocalic r, presents different realisations in different accents of English, some of them coinciding with the realisation of (r) in Spanish. Similarly, the voiceless postalveolar fricative/ʃ/, a phoneme present in the English phonetic inventory, occurs in Spanish as an allophone for the affricate /tʃ/.

conditioning factor (the earlier the speakers are exposed to the L2, the faster their production of L2 sounds). The age of acquisition was considered in the original design of the project, but given the homogeneity of the group (who were around eighteen years old at the beginning of their acquisitional process) and the similarity of the preliminary results, this factor was discarded in the analysis section; however, the late acquisition of the L2 may have an effect on, for instance, the fossilization of certain phonemes, as will examined in the discussion section (Section 5.2.4).

6. *Style*: Beebe (2006) examines the speech of nine Thai ESL speakers living in New York from three social classes; she studied two phonetic variables /l/ and /r/ in different word positions, in both conversational and listing style (less and more formal, respectively). These variables have a different behaviour in Thai: one of them is socially marked and the other not. The findings suggest that the use of either of these two variables depends on whether they are prestigious or not in the first language of the speakers, that is, if the variant is not socially marked, the L2 acts as the superordinate system, but if it is socially marked as prestigious in Thai, the L1 acts as the superordinate system; this variation depends on the formality of the interaction. As in my research, this study examines the behaviour of socially marked variables, therefore its results could shed light on the behaviour of prestigious versus non-prestigious phonetic variants.

2.1.3.2.1.2 Summary

By describing studies in immigrant context I attempt to show how non-native variation can be compared to native variation in situations where the learner has to deal not only with the acquiring language, but also with contextual information which may have an effect on said variation. These studies show that both internal and external constraints need to be considered in the examination of non-native English, thus providing theoretical support to the inclusion of the different factors selected in the design of my own research – but, will NNSs of English in immigrant/immersion situations exhibit the same patterns of variation as speakers in instructional settings? To answer that, I next review studies conducted in settings in which the L2 is not learned in direct contact / immigrant situations, but mostly in formal instruction settings, specifically in contexts where English is considered a foreign language (EFL).

2.1.3.2.2 Variation in foreign language settings

Following the focus of my research, studies suggest that the variation occurring in first (L1) and second (L2) language has also been found in foreign language acquisition contexts,[32] as shown by Regan (1996) in the study of French negation by Irish learners of French as a foreign language; or Major (1996) in the study of Brazilian Portuguese speakers learning English.

However, these kinds of studies are scarce when compared with other studies of variation in second language acquisition, or they are carried out using other methodological and theoretical frames from the variationist one. For instance, the relationship between production and perception of English vowels in Brazilian Portuguese has been studied (Rauber *et al.*, 2005; Barboza, 2007), and in Polish (Rojczyk, 2010), or the study of the substitution of flaps by Japanese learners of EFL (Riney *et al.*, 2000).[33] As with the previous section, in order to justify the inclusion of different factors in the analysis of my data, the results of variationist studies on foreign language situations are presented next.

2.1.3.2.2.1 Main findings for NNS in foreign language settings

This section introduces some relevant findings for the study of variation in NNS, particularly in instructional settings.

1. *Preceding and following segment*. This factor has been traditionally linked to variation in the study of phonetic features in native speakers of different languages (see e.g., Cedergren, 1973; Poplack, 1980; Hindle, 1980; Benson, 1988; Guy, 1994; Labov, 1994; Carlisle, 1997; Tagliamonte, 1998; Tagliamonte

[32] For practical purposes, I suggest to make a distinction between second language acquisition and foreign language acquisition: the first usually refers to the language acquisition processes that take place in natural settings such as immigrant – hereafter, this will be simply referred to as L2, whereas the second refers to those contexts in which the language is not the second language of the place in which it is being learned (Gass and Selinker, 2008: 268), and usually taught in formal instructions contexts – hereafter, as the focus of my research is in English, I will refer to this as EFL; thus, e.g., a L2 variationist study is the case of Alberto, a Spanish speaker learning English in the U.S. (Schumann, 1978) and a variationist study of EFL is the case of Brazilian speakers learning English in an institute in Rio de Janeiro (Major, 1996).

[33] These studies (Rauber *et al.*, 2005; Barboza, 2007; and Rojczyk, 2010) compare the acoustic differences between NS and NNS of English in EFL contexts. Studies like these represent an approximation to the study of variation in L2 contexts, but their methodology cannot be considered within a traditional sociolinguistic approach.

and Temple, 2005; Bayley, 1994; Santa Ana, 1996). One very relevant study for the present research is Major's 1996 study, which attempted to test whether universal patterns of phonetics and phonology are similar to the ones exhibited in the process of acquisition of a second language. He examined the speech of NSs of Brazilian Portuguese studying English in Rio de Janeiro, particularly in the production of consonant clusters. He classified the substitutions of phonemes in three categories: *transfer* (L1-like), *developmental*, and *correct* (target), alluding to the build-up process of acquisition, i.e., at the transfer level the speakers were producing L1 sounds, and at the correct level the speakers were producing target-like features (English). The results suggest that, in terms of sonority and markedness,[34] clusters that are more marked are acquired earlier than less marked clusters.[35] This means that most proficient speakers, producing target-like clusters, would produce more efficiently clusters that incorporate stops + fricatives than fricatives + liquids, i.e., the use of target-like forms will depend on the experience of the speaker using the L2, and that there are some features that are acquired first and others later, just as in first language acquisition. In sum, Major's findings suggest that the phonological environment is a crucial factor in the production of certain features, especially those that, being more alien to the L1 of the speakers, are more easily recognised and, therefore, produced.

In another study, Regan's (1996) analysis of the negation particle *ne* in French as a second language examined the effect of preceding and following segment in the rates of deletion of this particle in Irish speakers learning French before and after a brief period of immersion in France. The author found that there is a consistent difference in the effect of following consonants and vowels, in interaction with the immersion of speakers in France, with vowels favouring the deletion of *ne*; this trend was increased after the immersion period. He also assumes that the same hierarchies of deletion occur with NS, but the absence of contrasting data does not allow testing for this

[34] Markedness is defined by Hume (2011: 79-80) as a property of a given sound or linguistic element; in phonology, marked sounds (or groups of sounds) are opposed to unmarked sounds, as the latter can be considered "more frequent, natural, simple, and predictable" than the former. Thus, *marked* consonant clusters are made up of plosives+plosives consonants, and *unmarked* (or less marked) are made up of plosive + fricative or fricative + plosive consonants (see Hume, 2011 for a complete account on markedness).

[35] Considering as sonority scale vowels > glides > nasals > obstruents (Major, 1996: 97).

factor with NS. From these studies we learn that there is a consistent difference in the effect of consonants and vowels in NNS variation, which is also consistent with the literature for NS variation; therefore, I will include this as a potential constraining factor.

2. *Differences in the phonetic inventory of the interacting languages*. As with NNS learning of L2 in immersion settings, the absence of some phonemes in one of the two languages in interaction has also been studied in foreign language settings. Schmidt (1987) examined the two phonemes for *th-words*; this is /ð/ and /θ/ in Egyptian speakers of Arabic in relation to a series of factors, such as style, social class, and educational background. The hypothesis suggested that stylistic and educational backgrounds were responsible for the replacement of the English phonemes by non-native phonemes, but both groups (Classical Arabic literate and illiterate speakers) showed similar rates in the target-like use of the studied phonemes, testing that substitution is not the rule for most speakers, independent from their educational background. This particular factor is not examined as a constraining factor *per se*, but it is interesting to examine as part of the discussion on the potential origins of the variation of the selected variables.

In a different kind of study (lexical and morphosyntactic variation), Durham (2006) examines the variation of high-level[36] Italian, German and French NNSs of English in Switzerland, using English as a lingua franca. She aimed to test, first, if the three groups of speakers behaved similarly in their use of English, and second, if NNSs of English "are able to acquire the structured heterogeneity of the native English speaking community […] to determine if native variable patterns can be transmitted intact to non-native speakers" (Durham, 2006: 58-61). Her results showed three different patterns:

- Case A (variation acquired): the three groups of speakers show similar variation to native speakers of English in the use of relative pronouns (complementisers).

[36] Durham (2006: 59) defines high-level speakers (and not learners) as those who have "gone far beyond the initial acquisition stages and whose language can no longer be considered an interlanguage."

- Case B (problems of acquisition): the variation of the three groups of speakers is similar among them, but different to native speakers of English for the variable future (*will* versus *going to*)
- Case C (variation influenced by native language): the variation of the three groups of speakers was different among them and also different to native speakers for the variable non-count plural forms (*information* versus *informations*), and the adverbials *also*, *as well* and *too*.

Her conclusions indicate that "of the features considered [..., those associated with lexical variation appear to have been more difficult for non-native speakers to acquire than those in which the variation is tied to morphosyntax." The variation seems to be explained by frequency ("the lowest frequency features are the least well acquired" (37), and interference. Her study is relevant here as it shows that the variation of non-native speakers is likely to replicate native speakers', as an indicative of the acquisition of L2 features; however, it is also likely that, in absence of native-like variation, the linguistic variation of non-native speakers can be sourced in their L1 (in the form of transference).

The sections above show the theoretical framework that supports my study, They define variationist research from its principles and its methodology; the setting of the theoretical framework concludes with a review of the main studies conducted in variationist research for English and introduces the study of variation in L2 settings, both from a theoretical and empirical perspective. But, what are the characteristics of the L1 of the speakers (Chilean Spanish) from a variationist perspective? How can their linguistic performance in a L2 reflect their L1, as in Durham's abovementioned study? The next section introduces the study of variation in Chilean Spanish, describing the language from both a historical and linguistic view point.

2.1.4 The study of Chilean Spanish variation

The variation in Chilean Spanish is important to consider as (1) it will show that the use of certain features in native speaker speech is constrained by specific linguistic and extralinguistic factors, and (2) it will help to predict the behaviour of the speakers in their L1 in relation to the selected features. The following sections show that there are a number of studies which document the differences in Chilean

Spanish. Over the last 20 years, the development of sociolinguistic research on Chilean Spanish has not been as prominent as in other countries; however, a number of studies stand out: one of the most important works is Wagner's *Linguistic and Ethnographic Atlas of Chile by Regions* (1998, 1999, 2004), which describes the use of several features (at phonetic and morphosyntactic levels) by speakers of both urban and rural settings, of different age groups, and of all social and educational backgrounds.

Also, the speech of several cities has been registered and defined, identifying the characteristic features of Chilean Spanish. Some relevant findings in relation to the variables selected in my research indicate that, for instance (1) voiced dental fricative /ð/ is constrained by such factors as intervocalic position, morphological structure, sex, age, social class (Cepeda, 2001; Valdivieso, 1993; Cepeda and Poblete, 1993) and style (Pérez, 2007); (2) voiceless alveolar affricate /ʧ/ has shown at least three variants (Bernales, 1978; Valencia, 1993-1994) – also present in my research – constrained by sex, age, and social class (Bernales, 1978; Cepeda, 2001; Valdivieso, 1993; Widgorsky, 1978), and (intervocalic) position (Bernales, 1978; Widgorsky, 1978).[37] The subjective evaluation of variable speech has also been studied, showing that Chilean speakers are aware of the social meaning of particular features, especially in upper and middle class speakers with higher educational backgrounds (Valdivieso, 1981; Tassara, 1992). All these findings indicate that the patterns of variation of Chilean Spanish demonstrate ordered heterogeneity, and, therefore, the linguistic behaviour of Chilean speakers can be predicted on a quantitative basis. Thus, together with the patterns exhibited by NS of English, variationist research in Chilean Spanish introduces the patterns of variation for NS of Spanish. These studies set the basis for contrasting the patterns of variation of the speakers of my sample (L2) with the patterns exhibited in L1 contexts for both English and Spanish.

[37] I present a complete account of the main variationist findings for each variable studied in Chapter 4.

Lipski (1994) provides one of the most thorough characterisations of Chilean Spanish. This, conducted from the descriptive linguistics perspective, is based on the speech of Santiago and Valparaiso; this characterisation is, however, too broad, and does not consider the fact that between these two cities there is an acknowledged dialectal difference that has not yet been studied.[38] Valparaiso has been historically the main port of Chile, receiving the first waves of immigrants from different parts of the world, and, in the present day, thousands of tourists and cargo ships. In contrast, Santiago's population is relatively stable and highly stratified, with the less privileged classes concentrated in the north, west and south of the region, while the more privileged residents live in the east. The centre of the region, namely the capital, Santiago, is the area where most of the services are focused. However, there is a good reason to characterise the speech of all of Chile as being like that found in Santiago, as almost a third of the Chilean population lives in that area – the Metropolitan Region is inhabited by nearly six million people, and it represents the main economic, political, and educational centre.[39] Here, according to Wagner (1998), "differential features [of Chilean Spanish] have been neutralised by means of the formal variety distributed along the country", suggesting that variation exists through an informal variety which is used in the four different dialectal zones. These

Figure 2.2 Map of Chilean dialectal zones

[38] Some relevant studies in Chilean sociolinguistics carried out in Santiago are: *El estudio del habla culta de Santiago de Chile* (1967-1983) by Rabanales and Contreras (1990) and *Las muletillas en el habla culta de Santiago de Chile* by Contreras and Rabanales (1991). Lidia Contreras has also worked on the inclusion of borrowings in educated people (1988). This will be addressed in the Chilean sociolinguistics section (Section 2.1.4).

[39] Data taken from the 2002 Census.

four areas have been delimited in the *Linguistic and Ethnographic Atlas of Chile by Regions* (Wagner, 1998, 1999, 2004).

At a phonetic level, Lipski (1994: 197) describes the following features as the most prominent in Chilean Spanish:

- The aspiration and elision of /s/, as in /a'ßlamos/ > [a'ßlamoh] or [a'ßlamoØ]
- The realisation of /n/ as alveolar in word-final position in most Chile, but as a velar in northern Chile. /pɪ'ɲon/ > [pɪ'ɲoŋ]
- Neutralisation of /ʎ/ and /y/, favouring a palatal fricative [y] /'vaʎja/ > ['vayja]
- The neutralisation of liquids, which results in the assimilation of /rn/ > [nn], as in /'karne/ > ['kanne], /rl/ > [ll] as in /arle'kin/ > [alle'kin]
- The assibilation of the trill /r/, the dropping of word final /r/ or its realisation as a partial or total voiceless sibilant /tra'ßaxo/ > [tˇɹa'ßaxo]; /a'ɾar/> [a'ɾaØ]> [a'ɾaˇɹ]
- The fricativisation of the affricate /ʧ/, as in /ʧile/ > [ʃile], or realised frequently as a prepalatal, almost a [ts]
- The assibilation of the group /tr/, as in /tra'ßaxo/ > [ʃɹa'ßaxo]
- The realisation of the voiceless velar fricative /x/, as voiceless palatal fricative [ç] before front vowels /ʌ'xɪ/ > [ʌ'çɪ]
- The devoicing of unstressed word-final vowels /'gato/ > ['gat̥o]

Other phenomena registered in Chilean Spanish are the ones described below by authors such as Cartagena (2002), Cepeda (2001), Contreras (2002), Salamanca and Marlett (2001), Haensch (2001), and Saez (1999):

- The (post) palatalisation of velars, as in /soga/ > [soɣa]
- The fricativisation or elision of /d/, as in /dedo/ > [deðo] or [deØo]
- The devoicing of /β/ and /ɣ/ /se' βaðo/ > [se'ɸaðo]; /'raɣa/ > ['raxa]
- The bi-labialisation of the labiodental /f/ as in /a'fuera/ > [a'ɸuera]
- The change from hiatus to diphthongs in /ea/ > [ja] as in /reak'sion/ > [rjak'sion], /oa/ > [wa] as in /to'aʎa/ > [tw'aʎa]
- The vocalisation of the bilabial in the group /bl/ as in /a'blar/ > [au'lar]

- The prosthesis of /g/ before the diphthongs /we/, as in /'weko/ > ['gweko] and /wa/ as in /'waso/ > ['gwaso]

Several of these phenomena are important for the present research, in particular the fricativisation and palatisation of the affricate /ʧ/, the fricativisation of /d/, and the multiple realisations or /r/, as these constitute the core features under study here. Despite the fact that my study examines the variation of Chilean speakers in English, the description of the variation found in Chilean Spanish is highly relevant, as one of the aims of my study is to test for transfer of sociolinguistic patterns from the L1 to the L2 and another is to test whether the variation exhibited by NS is similar to NNS of English.

But how is variation in Chilean Spanish justified from a historical perspective? This is how the patterns of variation can be explained in terms of e.g., the influences from other languages or dialects of Spanish. The following section presents a brief account of Chile as a linguistic body, with an emphasis on the possible effects of other languages through the diverse migration processes that occurred in the country.

2.1.4.1 Chile as a linguistic body

Chile has been described by many linguists as a very special dialectal zone. The intonation, lexical and phonetic inventory differ greatly from the other Latin American Spanish dialects. The natural borders, such as the Atacama Desert in the north, the Andes mountain range in the east and the Pacific Ocean in the west, have isolated the country, making it a cultural, social, economic and linguistic island. Nevertheless, Chilean Spanish has been shown to be a permeable language, subject to influences from other languages and varieties of Spanish, especially at the phonetic level. In that line, the final section of this chapter reviews the main migrant influences in Chile, which are thought to contribute to the linguistic development that characterises the Chilean Spanish dialect (e.g. Rabanales, 2000). This is done in an attempt to show how Chilean Spanish has adopted foreign elements, modifying its phonetic and lexical repertoire.

2.1.4.1.1 Immigration

As in other countries of Latin America, Chile has received immigrants from different parts of the globe. Previous to the Great War, at the end of nineteenth century, the saltpetre industry in the north of Chile brought many British, German and North American businessmen, who were attracted by the great earnings produced from the exportation of the mineral to different parts of the globe; originally from these languages we can find in Chilean Spanish words such as *jumper* (although in Chilean Spanish this refers to a female school uniform), *huachiman* (watchman) and *guaipe* (wipe) from English, *kinder (*Kindergarten*)*, *schop* (draught beer), and *strudel* from German. During this period of time, nearly 2000 Chinese workmen also arrived to work in the saltpetre mines, and their language gave to Chilean Spanish the words *té* (tea) and *charol* (shiny varnished material used mainly in shoes and bags). Not many studies have been carried out on the influence of foreign languages on Chilean Spanish; Rabanales (2000: 138), for instance, points out that Chilean Spanish is made up of five groups of lexemes: peninsular (referring to Spanish spoken in Spain), creole, indigenous, mixed, and foreign,[40] suggesting that the contact between natives, Spanish, and immigrants has been essential in the formation of Chilean Spanish as a dialect (in lexicographic terms), which keeps changing and evolving, especially in informal speech. All the examples above confirm that Chilean Spanish has historically received foreign influences, at least at the lexical level. Recently, Saez (2005) has written on the presence of Anglicisms in Chile, introducing the function of the Anglicisms on Chilean Spanish, the degree of adaptation to Chilean Spanish, the phonetics (use and adaptation), the morphosyntactic behaviour and the semantic use of the Anglicisms, the ability of Spanish to absorb and react to them, and the attitude of the speakers towards them, amongst others topics.[41] Dieguez (2004)[42] also reports on the use of Anglicisms in the economic discourse of scientific news articles; based on translation theory, she describes the use of the Anglicisms in contrast to their equivalent in Spanish in relation to factors such as discourse modality, speaker,

[40] He distinguishes two groups: the first made up of words that were already adopted by the Spanish at the time of colonization from languages such as Greek, Arabic, Hebrew, Persian, Germanic, French, Italian, Chinese, Japanese, and Turkish; and a second group of more recent incorporations that include the languages already mentioned, but especially from English (Rabanales, 2000: 139-140).

[41] This article is a descriptive view on Anglicisms in Chile, not a quantitative study that shows the use or the presence of Anglicisms in different domains of speech.

[42] As in Saez (2005), this is also a descriptive and functional work, not a quantitative one.

audience, communicative intention and space-time situation. From these studies, we learn that Chilean Spanish is in fact an accesible language, capable of receiving new sounds and words from other languages, not only using them as borrowings (in their original form such as *factoring*, *duty free*, *outlet*) but also incorporating them with phonetic and graphemic adaptations to Spanish (as in *guaipe* for the English *wipe,* or in the form [ʧəʊ] for show). This is important for my own study, as it shows that Chilean Spanish not only uses non-native features (at the phonetic level), but also adapts them to the local use, therefore, making its phonetic inventory permeable to non-native sounds, including English-like phonemes and making English potentially easier to acquire. If foreign features influence Chilean Spanish, how much of Chilean Spanish is taken into English? In terms of the effects of the inclusion of these foreign elements, especially of those which have been already adapted, a closer use to the Spanish variants could be expected, particularly in relation to cognate elements and to those phonemes that are present in the phonetic inventory of both languages.

2.1.5 Summary

The orderly heterogeneity of language variation, as proposed by Weinreich, *et al.* (1968) has been shown to be present in first and second language situations, constrained by internal and external factors. The above sections have described these, reviewing the theoretical foundations for the study of variation and presenting the main findings for the variation exhibited by native and non-native speakers of English and Spanish. The next chapter presents all the methodological considerations taken in the development of the present research, from the design of the study to the definition of the analysis techniques, which will later be used in the actual analysis of the data obtained.

CHAPTER 3 - DATA AND METHODOLOGY

The purpose of this chapter is to present all methodological aspects related to the study of variation, specifically in relation to data, from the design of the actual research to the data collection and analysis. The first sections introduce the research questions and hypotheses, which are followed by the sample selection, ethical considerations, data collection - both in design and in practice - then end with the manipulation (transcription, coding) and analysis of the data.

3.1 Introduction, research design: research questions and hypotheses

The purpose of all variationist research is the analysis of the actual use of language by a particular group or community of speakers. The process of collecting the appropriate data implies several processes, such as the identification of the objectives of the research (what is the purpose of examining the speech of Chilean speakers of English?), the definition and selection of the sample (what data do I need and who will provide me with the best data to satisfy the objectives of the research?), and the design of the interview (how will I elicit the data I need?) (for more on data collection, sample definition, and elaboration of the interview see e.g., Labov, 1984; Tagliamonte, 2006). The following sections will address these issues.

3.1.1 Research questions

One of the objectives of any variationist enquiry is to answer the research questions initially identified in the design of the research itself. From the arguments presented in Section 1.1 (*Aims for the present study*), I formulate the following research questions:

- Is phonetic variation systematic in Chilean speakers of EFL? If so, is it attributable to social and/or linguistic factors?
- Is the variation similar to that shown by NS of English in relation to the same variables? If not, how do they differ?

To answer these, the appropriate data needs to be obtained; thus, what are the criteria that help the researcher identify the best speakers that will be part of their sample? What have the previous researchers looked at? Milroy and Gordon (2003) suggest that representativeness is a key issue to be considered. The representation of a sample aims to include all relevant characteristics of a population in the sampling universe; therefore, it is essential that the design of the research considers all the characteristics that the researcher is interested in studying in a specific group. In the case of my research, to test for the systematicity of variation in Chilean speakers of EFL the most important prerequisite that needed to be met was that participants were Chilean speakers of English (the other criteria used for the selection of speakers is provided in Section 3.2.1.2.2, *The Informants*).

Also, based on the previous research on the variables described in Section 1.2 (*Introduction to the variables*) and on the principles of variationist research described in Section 2.1.1. (*Principles of variationist research*), I formulate the following hypotheses.

3.1.2 Hypotheses

The rationale behind the formulation of the hypotheses has basically two origins: the most common principles of linguistic variation drawn from the relevant literature and the observed results from studies of NS and NNS of English. I also distinguish two groups of hypotheses:

Primary

- Variation in Chilean Spanish speakers of English is systematic.
- Hypercorrection occurs as a result of the transfer of prestigious variants from Chilean Spanish to English.
- Males and females pattern differently.[43]
- The amount of exposure to English in formal education is equivalent to the use of target forms, i.e., final-year informants use more target forms than first-year informants.

[43] As this is not a case of stable nor changing linguistic situations, we cannot contrast the claims made, e.g., by Labov (1990) in relation to the gender paradox (see Section 2.1.2)

- The phonetic environment is a relevant contributing factor in variation.

Secondary

I will develop specific hypotheses for each of the four variables – voiceless alveolar fricative (ʃ), voiceless alveolar affricate (tʃ), voiced dental fricative (ð) and postvocalic (r), based on previous studies on the selected variables, plus the general hypotheses for NSs and NNSs of English. These are presented in Chapter 4.

Now, how do I answer these questions, and how to test for the formulated hypotheses? Is the selection of the sample determinant to conducting the analysis of the chosen features? Without the appropriate sample, the study makes no sense, as it is designed to test for specific hypotheses. In Section 3.2 below, I describe the process from the articulation of the hypotheses to the definition of the sample.

3.2 From the research questions to the sample

To fulfil what has been described as minimum requirements in terms of representativeness (see Tagliamonte, 2006), the following sections attempt to describe how the sample was selected (based on the formulated hypotheses) in terms of its requirements and the origin of the individuals, as well as presenting some general demographic characteristics.

3.2.1 Requirements to define the sample

Judgement sampling was used for this research, in which "the researcher identifies in advance the types of speakers to be studied and then seeks out a quota of speakers who fit the specified categories" (Milroy and Gordon, 2003: 30, see also Tagliamonte, 2006). I followed a number of criteria in the selection process.

Tagliamonte (2006: 23) suggests that "a minimum requirement for any sample is that it have a degree of representativeness on the basis of age, sex and (some way of determining) social class, education level or both". In terms of these usual social dimensions (see e.g., Milroy and Gordon, 2003; Bayley and Lucas, 2007; Labov, 2001; Tagliamonte, 2006), sex and educational level were incorporated as requirements.

The inclusion of sex has traditionally been used to reveal distinct variation between male and females. While most research defines the level of education in years (see e.g., Meechan and Foley, 1994; Tagliamonte, 1998), educational level is considered differently here, as my study is restricted to university students in different stages of their programme: initial (first year), intermediate (third year) and final (fifth year).[44] If this were a case of L1 research, and with such a short span between the different educational levels of my sample, the differences in variation might not be really significant, but as this is a case in which learners are being exposed to language progressively, it is expected that these differences are more meaningful.

The inclusion of age as a factor had no particular motivation, as all speakers were intended to belong to the same age group, that is, all university students 18-25 years old; along with age, social class was not considered a selection criteria, as social class is a highly sensitive topic so its inclusion would have a negative effect on recruiting informants (for more on the factors considered see Section 3.4.3.2 below). Other aspects considered in the characterisation of the sample (educational background,[45] immersion in L2 settings, etc.) are presented in Sections 3.2.1.1.1 and 3.2.1.2 on the definition of the sample.

3.2.1.1 The community

A community, in Labovian sociolinguistics, is made up of individuals who adopt and recognise the norms of language use for a specific area or group. In my research, this sort of speech community consists of learners of English as a foreign language. The first step is, then, to define the binding factor for this specific community of speakers: acquirers of English in a foreign context.

3.2.1.1.1 The teaching of English in Chile

At the beginning of the last decade, the Chilean Government launched a programme with the objective of making of Chile an English-Spanish bilingual country. Some of

[44] In terms of years, students in first year have been educated formally for at least twelve years, in third year, fourteen years, and in fifth year sixteen.

[45] I consider educational background as different from educational level. The first distinguishes the type of education of the learners (a public school it is not the same in terms of quality as a private school), and the latter the amount of time learners have been in formal education.

the policies included increasing the hours of English in schools, a change in methodology – from a grammatical approach to a communicative one, the increase of university programmes dedicated to training teachers, and many international treaties with English-speaking universities and institutions, such as the British Council. This section has a double aim: first, to present English in primary and secondary educational settings, as this is where the individuals of the sample come from and where most of them will eventually teaching, and second, to introduce the context where the learners acquire EFL. The last part of this section is dedicated to the description of the accent of instruction, Received Pronunciation (RP).

3.2.1.1.1.1 The teaching of English in primary and secondary education

Currently, the teaching of English in primary and secondary education depends on the kind of school the students attend, which is socioeconomically bounded: private and some subsidised schools range from four hours a week of English language as a subject, to all their subjects taught in English; some private schools (the most prestigious and those with the highest fees) even have native speakers of English on their staff (most teachers of English are non-native speakers, trained in local universities). Public schools range from two to four hours of English lessons a week, in which most classes are characterised by an overpopulation of students, poor access to classroom materials, and little emphasis on the communicative approach sought by the national curriculum.[46] The aim of the national curriculum is to reach B1 level (PET – Preliminary English Test) of the Common European Framework of Reference for languages (CEFR), but results show that most students are below A1, that is, they do not achieve the basic competences required to communicate in English (Ministerio de Educacion, 2012)

Furthermore, studies linked to the Ministry of Education show that teachers at primary and secondary levels are scarcely certified, some of them not meeting the requirements suggested by the Ministry of Education (2012) (B2 level, Vantage). This means two things: that the students leave secondary education with less than the

[46] As a teacher, I taught English in a rural school that used the curriculum designed by the Ministry of Education, and the characteristics of the curriculum are not applicable to it, as with the vast majority of public schools with few resources, overpopulation of students, low number of hours, etc. My research does not focus on the process of learning English primarily in primary and secondary educational settings, but on the variation exhibited by university students; therefore, I will not discuss the quality of the Chilean educational programme, or similar topics.

minimum skills to communicate in English and probably, as an extension of this, that the teaching of English is not being effective, and second, that teacher training is not successful, as most of the teachers are below the suggested minimum requirements for teaching.[47] But, how are the teaching programmes designed and how does the acquisitional process takes place? These two questions are answered next.

3.2.1.1.1.2 The teaching of English in Universidad del Bío-Bío

In the case of the speakers of the sample, they are students on an English teaching programme to become teachers of English, specifically in secondary education.[48] The English teaching programmes vary in the different universities and institutes that prepare teachers: from four to five years, with lessons for beginners from day one, increasing progressively in difficulty (intermediate, advanced). The subjects include English culture and literature, along with educational methods and theory, educational psychology, general linguistics, etc. The access to traditional universities is via a national admission test,[49] in which students with higher scores access the most prestigious programmes (medicine, engineering, architecture, and the like) and the most universities.[50]

In more technical terms, UBB is one of the most important universities in Chile; it ranks 15 in Chile (out of the 19 universities ranked) and 158 in Latin America (out of the 192 ranked universities) (SCIMAGO, 2014). Its programme, "English Teaching",[51] has prestige in the national English teaching community and its curriculum has been accredited by the National Committee of Accreditation, an organisation in charge of

[47] The Chilean Ministry of Education (2002) has set B2 as the minimum level of English required for teachers. The percentages of certification for teachers are: Pre A1, 1%; A1, 3%; A2, 9%; B1, 21%; B2, 37%; C1, 23%; C2, 6%. The levels A1 and A2 correspond to elementary, B1 and B2 to independent users, and C1 and C2 to competent users, according to the CEFR.

[48] Very few programmes in Chile prepare teachers for elementary education; in fact, most teachers of English in primary education are primary teachers who have undertaken English courses.

[49] Private universities set the requirement of sitting for the exam, but they do not consider the final score; their students usually pay considerable amounts of money in tuition fees.

[50] Given the highly stratified educational system in Chile, the best schools, which require fees, produce the highest scores and their students can access the most prestigious programmes; on the other hand, less privileged individuals only have access to public, free schools, in which the quality of the education has been recently under scrutiny; these public schools produce the lowest scores and their students can only access less prestigious programmes and universities, thus reproducing the social class model.

[51] At the time of the interviews, the fifth year of the programme was the last generation of the "English Teaching and Translating Programme", as, thereafter, changes in the design of the programme resulted in the translation area being removed.

assessing the quality of the universities and their programmes. The admission score average is 600 points (admission scores range from 850 to 450; any score below 450 does not allow the candidate to complete the admission process).

Traditionally in Chile, teaching programmes do not attract the highest scores (as science-related programmes do), and, therefore, the scores for this programme locate it among the highest in this area in the national context. As a result, students from different parts of the country study English teaching in UBB, making it a rather plural community.

The English Teaching Programme is a five-year programme, with each year divided into two semesters in which students must attend, and be approved in,[52] several subjects, such as pre-intermediate, intermediate, post-intermediate and advanced English language, phonetics and phonology, English grammar, Anglo-Saxon culture and civilisation, English literature (several periods), drama, linguistics, as well as pedagogical subjects, psychology, and quantitative educational research. From the fifth semester, students start attending public, private and subsidised schools as part of their teaching experience, and in the final year they must produce a small research project in any of the areas covered during the programme.

In relation to the English language area, students face the language from their first day for about 16-24 hours a week; by design, this programme attempts to produce bilingual, native-like speakers,[53] or, at the least, proficient users of English; therefore, the interaction with the L2 is essential: the more they are exposed to the L2, the closer they are expected to perform to a native speaker. This is highly relevant for my research, as one of the objectives, as stated in the research questions, is to determine if the variation exhibited by NNSs is comparable to NSs. At the time of the interviews (December 2006), only two of the six teachers were native speakers of English (from England and Canada), and all the other teachers were Chilean and educated in Chilean universities. This means that the variety of language that the speakers of the sample are exposed to is mainly characterised by 1) materials designed for ESL purposes, i.e., recordings or written texts produced by native speakers of English (mostly RP

[52] If students fail a subject, they are not allowed to move forward and they must repeat the subject.

[53] By 'native-like individuals', I refer to those who perform similarly to native speakers in relation to their use of target or native forms at different language levels (phonetic, syntactic, and discursive).

speakers) in both artificial and real-life situations and 2) a spoken face-to-face variety of English produced by NNS. Potentially, this may have an effect on the output of students, if we consider this a mirror of the input; however, given that the gathering of the data did not consider or measure the sources of input as conditioning factor, it is not possible to assert this. The accent of instruction varies from institution to institution. For instance, since 2012 I have worked in two places: from March – December 2012, I worked in Instituto Chileno-Britanico de Cultura de Valparaiso, where most classroom material is designed by Cambridge University Press (this is also a Cambridge ESOL Examination centre), with a strong emphasis in Contemporary Bristish English, and during the period July 2012 – July 2014, I worked in the teaching programme of Pontificia Universidad Catolica de Valparaiso, in Vina del Mar, Chile, where the syllabus does not refer to any particular accent or dialect of instruction; furthermore, most of the academic staff are non-native speakers of English (Chilean) who speak with no particular accent, i.e., with traces of GA or RP, but not as marked so as to define them as GA, RP or any other English accent or dialect.

Most of the English language classes in UBB are based on the use of textbooks and recordings; as they progress, the textbooks and recordings are replaced by conversation-based tasks, topic-based presentations, and discussions. The classes are made up of about 60 students in the first year and the numbers decrease progressively due to various circumstances, such as desertion or as students fail to pass their subjects,[54] i.e., a typical third year is made up of 40 students and a fifth year of about 12 students.

Having introduced the general characteristics of the university and the English teaching programme, now I focus on the accent of instruction used in many SLA classes, including the one from which the sample individuals were selected.

3.2.1.1.2 The accent of instruction (RP)

According to Hughes, *et al.* (2005: 2-3), RP (Received Pronunciation) is the most typical accent taught to foreigners. Originally from London and its surrounding areas,

[54] As mentioned above, this programme is made up of five levels, each one of which students must pass to move to the next level; so, as some move forward, others must repeat the same course, increasing the number of students in the class. Also, not all students continue their education, leaving their studies unfinished; this is more common on higher-level courses.

RP is regarded as the most prestigious accent, usually heard in British public schools and in the "upper reaches of the social scale" (*ibid*.). Hughes, *et al.* (*ibid*.) suggest three main reasons why RP is the most learned accent: firstly, given its high prestige, it is considered "the best, the clearest and even the most 'beautiful' accent"; secondly, this accent is traditionally linked to BBC newsreaders, therefore it is recognised worldwide as easily understandable, so the successful learners of RP will have better chances of being equally (easily) understood; and thirdly, it is the most described accent in literature and, consequently, information about RP is easily accesible for non-native speakers in instructional settings (see e.g., Jones, 1917; Wells, 1982; Gimson, 1988; Cruttenden, 1994; Ladefoged and Maddieson, 1996; Milroy, 2001; Hughes *et al*., 2005).

As one of the preferred accents of instruction in many SLA contexts (namely, institutions),[55] RP is also used in primary and secondary education as a mirror of the accent acquired by the teachers in the different universities; most of the universities seem to base their programmes on the teaching of RP using texts such as *Gimson's Pronunciation of English*, and textbooks from Cambridge, Longman and Oxford publishers (e.g., the *Interchange* series from Cambridge University Press, *Wells' Longman Pronunciation Dictionary*, *Oxford Advanced Learner's Dictionary*, and the like).[56] In the case of UBB, the curriculum, at the time of the interview, was based on non-authentic material that was specifically designed for EFL purposes, especially in the first years. As the students progressed, they were able to use authentic pieces of writing (e.g., literature, newspapers, etc.) and audio (interviews). Contact with native speakers (particularly with RP speakers), was very limited, given that the country is physically far from English-speaking countries and also because the possibilities of exchange are very limited, economically speaking. As a consequence, most speakers who leave UBB rarely have experienced immersion situations or have had sustained prolonged interaction with native speakers of English.

[55] More recently, RP shares the status of preferred accent with General American (GA); GA is examined in the results section, in connection with one of the four variables selected; see e.g. Hughes *et al*. 2005.

[56] From the time of the interviews to the present, the situation is perceived as different. Most institutions are using authentic material from the beginning of the instruction process for two reasons: first, it has been acknowledged that the status of RP has changed over time (Hughes *et al.* 2005), and second, providing students with real written and oral text allows them to receive a wider spectrum of accents, which in turn would allow them to recognise and use English in a wider variety of contexts.

In relation to the characteristics of RP, this is described widely as a non-rhotic accent, where the /r/ that follows a vowel is not produced, as in *farm* [fɑːm]; at segmental level, it is characterised by nineteen stressed vowels (/ɪ, ɛ, a, ɑ, ʌ, ʊ, ɛː, iː, ɑː, ɔː, uː əː eɪ, ɔɪ, ʌɪ, aʊ, əʊ, ɪə, ʊə) and two unstressed vowels (i, ə) (Upton, 2008: 240); in relation to the consonants, it shares the same phonetic repertoire with other varieties of English (*ibid.*); however, some differences are found, for instance, in the articulation of /f/ and /θ/ in British and American English (apical-dental and interdental, respectively; Ladefoged and Maddieson, 1996). These consonantic differences are widened at the phonemic level (see e.g., Wells, 1982; Hughes *et al*. 2005; Ladefoged and Maddieson, 1996).

RP is characterised by phenomena such as glottalisation (associated with /t/ in "syllable-final position preceding a non-syllabic consonant as in *rat trap"*, and not in intervocalic position, as in other varieties of English, Upton 2008: 249), linking ([fɑːr ʔəˈweɪ]) and intrusive /r/ ([lɔːr ənd'ɔːdə]), yod coalescence (/tj/, /dj/, /sj/, and /zj/, /dʃ/, /dʒ/, /ʒ/) and deletion (as in the first syllable of super /sjuː-/ - [suː]), among others (Upton, 2008). For an extensive review on RP, see Jones, 1917; Wells, 1982; Gimson, 1988; Cruttenden, 1994; Ladefoged and Maddieson, 1996; Milroy, 2001; and Hughes *et al.*, 2005, amongst others.

As with all accents, RP is also subject to variation. Hughes *et al.* (2005: 2-11) describe the variation of RP at different levels. For instance, in terms of pronunciation, phenomena such as the simplification of triphthongs and diphthongs into single vowels, and the use of rising intonation in questions are frequent in RP speakers. Also, stylistic differences have been found in phenomena such as elision, vowel weakening and assimilation (*ibid*). Finally, Hughes *et al.* report on the effect of regional accent on RP speakers, arguing that RP is not a regional variety and that its speakers may be from a different social origin or region, moving along the social scale and modifying their accent towards RP.

In conclusion, RP is chosen by many learners and NS of English for many reasons, which include its prestige and its capacity to be understood internationally by other speakers of English, both NSs and NNs. The use of particular phonemes of RP is examined in relation to the four variables selected, and their use and variation is presented in the results chapter (Chapter 4).

Having defined the general characteristics of the community of speakers, the following sections continues with an actual description of the speakers of the sample, which is made up of eighteen university students from different educational and social backgrounds. These speakers belong to approximately the same age group, given that they are all undergraduates. The bond of this community is, therefore, the acquisitional process of English as a foreign language for instructional purposes. The rationale to select these speakers is also justified in the following sections.

3.2.1.2 The sample

3.2.1.2.1 Contacting the interviewees

I directly contacted some of the interviewees, some were contacted by the staff of the English teaching programme, some responded to advertisements posted in the English Teaching Programme Department at UBB calling for the interviews,[57] and others were contacted via interviewees, in what it is defined as *social network approach* (Tagliamonte, 2006: 21-22), particularly the non-local speakers, since some of the interviews were conducted after the academic term had finished and they were in their respective place of origin (Santiago).

The subjects were asked to participate in my project, which was described as doctoral research on linguistics; as the term "social" is commonly associated with social class – a sensitive topic in Chile – I did not use the term *sociolinguistics* to prevent a negative impact or unnatural behaviour influencing the performance of the speakers. Furthermore, as this term is not found in common conversation, another reason to avoid it is that it might lead speakers to believe that the conversation might be beyond their language skills' scope.

[57] The posted advertisements called for: "[…] students of first, third and fifth year to participate in an interesting study on the speech of students of the English Teaching and Translating programme." The details of the recording and the aims of the project (objectives, who would get access to the recordings, how the results would be published, etc.) were delivered personally to the interested subjects. These are included in Appendix B.

3.2.1.2.2 The informants

Most of the students are from Chillán, (407 kilometres south of Santiago) and its surroundings.[58] Chillán is a small city of about 170,000 inhabitants and the main economic activities are agriculture and public services. The profile of students that enter Universidad del Bío-Bío corresponds to the lower-middle classes, especially in educational programmes, which might result in a fairly homogenous sample group.[59] From these two characteristics, only social class will be considered in this study, given that, as mentioned earlier in relation to sociolinguistic studies carried out in Chile, no research has been conducted on the speech of local individuals (Chillán) so that locality could be contrasted with, e.g., Santiago, and, if it were the case, the study variation would have been in Chilean Spanish and not in English.

As a result of the judgement sampling, the criteria to select informants were: sex, with an equal number of males and females and years of instruction in English, with equally distributed informants from first, third, and fifth years (so progression in the instructional process is represented equally). Thus, nine female and nine male speakers, six from each academic year, were selected. Table 3.1 shows the speakers from the sample arranged by sex (male and female), and years of instruction in English (YIE – 1, 3 or 5).

[58] Refer to Section 2.1.4 The study of Chilean Spanish variation.

[59] Refer to the section on the teaching of English in Chile for more details on the social stratification of the Chilean educational system (pp. 63-67).

Table 3.1 Speakers arranged by sex, years of instruction in English and age			
Speaker	Sex	YIE	Age[60]
Paola	F	1	20
Cecilia	F	1	19
Marcela	F	1	21
Consuelo	F	3	20
Caterina	F	3	21
Josefina	F	3	23
Melinda	F	5	23
Carola	F	5	26
Carmen	F	5	27
Francisco	M	1	20
Jaime	M	1	22
Sebastian	M	1	21
Bernardo	M	3	22
David	M	3	24
Juan	M	3	20
Cesar	M	5	23
Andres	M	5	30
Miguel	M	5	24

As Table 3.1 shows, e.g., Paola[61] is a female speaker from the first year of the English teaching programme, aged 20. As age is not considered a conditioning factor, it is displayed in this table only for illustrative purposes to show that students belong to the same age group.

During the selection process and during the interviews, other demographic data was obtained from the interviewee data sheet, e.g., information on the level of education and occupation of parents, educational background of informants (private, subsidised, or public school), previous instruction in English (that is, in places other than university), and whether they had travelled abroad. From the interviews, different sorts of information arose: interaction with native speakers, differences in accent preferences, experiences with English language in media, etc. From these two groups, the demographic data sheet and the actual interview, only the first one was included in the social class indexation (see Section 3.4.3.2.3 for more on the socioeconomic group definition and indexation). The second, the interview itself, although it provided useful information about, e.g., interaction with native speakers or accent preferences, provides uneven and different sorts of data from the

[60] At the time of the interviews, end of December, 2006.

[61] As it will be detailed in the section on ethical considerations (3.3.2.1), the names of the speakers were replaced with pseudonyms in order to protect their identities.

informants and it is very difficult to assign a quantitative (even qualitative) value to the implications of such information. In other words, within the design of the present study, it cannot be tested whether informants' preference of British English over American English had any effect on their actual performance (except, probably, if their preference for British English results in a higher use of RP features, but again, this is not the focus of my study).

From the group of speakers, only one of them had received formal instruction in English besides university: Josefina (third year). At the time of the interview, one speaker had already (recently) left the programme (Carola), but she did not continue formal instruction in the university, or elsewhere, so, in educational terms, she is not different from the other fifth-year speakers.[62] The other speakers in the sample do not display any characteristic that makes them stand out from the group, which results in a fairly homogenous sample group.

Having selected the sample based on the research questions I attempt to answer, I next proceed to the actual collection of data, which is described in the following section.

3.3 Collection of data

3.3.1 Introduction

The design of the research requires the elicitation of natural conversation as this provides the most systematic speech, for which the traditional method used in sociolinguistics has been, since the 1960s, the interview. Labov (1984) suggests that the vernacular is the ideal style, as this is where speakers pay less attention to their language performance; however, with the sample and design of this research, it is not possible to obtain vernacular speech based on the definition given by Labov ("the mode of speech that is acquired in pre-adolescent years" (1984: 29). Most of the students from this sample began their acquisitional process when they were 18 years old; also, the natural setting of language acquisition is not the case for the speakers of

[62] Fifth-year speakers finished their studies in December, at the same time the interviews were conducted. In fact, all of the informants had finished their corresponding level by the time of the interviews, as the academic year starts in March and ends in December.

this sample); in EFL situations, which imply formal instruction, we might expect, at best, a casual style.

The sociolinguistic interview is a common approach defined by Labov (1984: 32) as "a well-developed strategy" designed to respond to a number of goals which will provide the researcher with reliable data; thus, this is the method selected to obtain the data with this group of speakers. Other methods, such as the reading of passages or of word lists, were discarded, as they are primarily used to test for variation due to changes in style (for a complete account on style, see e.g., Labov, 1972, 2000; Bell, 1984; Eckert and Rickford, 2001). Following the objectives of the interview proposed by Labov (1984: 32), in the following sections I outline the interview, from the process of collecting demographic data to the conversational modules and the actual design of the questionnaires that would be applied to the interviewees.

3.3.2 Interview preparation

Before the interviews took place, a written form with demographic data was filled in by each participant (see Appendix C). The form filled in by speakers contained details such as place and date of birth (to determine age and origin), the kind of school in which they completed their elementary and secondary education (an indicator of whether they attended private, subsidised, or public school – a social marker in Chilean society), and the level of education and occupation of their parents (another marker in the social indexation); they were also asked whether or not they had travelled abroad to indicate: first, if they had had a real experience in an English-speaking country, and second, their purchase power (another social marker).

The questions in the forms respond to the necessity of elaborating the socioeconomic index, as one was used by Labov in the New York study (cf. Labov, 1966, 1968, 1972 for more on socioeconomic indexation). Although Labov used a modified version of the Mobilization for Youth questionnaire (see Labov, 1972: 112-116) – in which the occupation of the breadwinner, the education of the informant and the income of the family group is equally weighted – I consider all of these elements and introduce them in the description of the socioeconomic groups without considering any statistical methodology, i.e., its elaboration was subjective, based on the information provided by the speakers (for more on the elaboration of the socioeconomic index,

refer to Section 3.4.3.2.3). This information provided in the form complemented the interview itself, considering that one of the ten goals of a sociolinguistic interview, as defined by Labov (1984: 32), is the collection of "demographic data necessary for the analysis of sociolinguistic patterns"; in other words, the interview will reveal in greater detail the general background of the speaker, particularly in relation to those aspects that are sensitive to ask the speaker directly, such as social class.

3.3.2.1 Ethical considerations

In accordance with the Ethics Policy of the University of Glasgow, ethical approval for the research was given by the College of Arts Research Ethics Committee. As part of Ethics Policy, before every interview started a consent form authorising the recording and the use of personal data for research purposes only was signed by every speaker (see the form in Appendix A). The consent form indicated that the interviewees authorised the use of the obtained information only by the researcher, and by third parties involved in the research (supervisors) with the prior authorisation of the researcher. The consent form also stated that their names would be replaced by pseudonyms so that their identity would be protected throughout the project. Furthermore, it was specifically stated that the recording of the interview would not be covert, despite the implications of this for the observer's paradox (Labov, 1972: 209).[63] The forms were written in Spanish and English, so that the individuals were fully aware of the nature of the interviews and how the data would be secured.

3.3.2.2 Questionnaire

The interview is based on two pre-designed questionnaires (see Appendix D), one for first and third year students (34 questions) and the other for fifth year students (33 questions). The use of a questionnaire is justified as it provides the researcher the ability, following the design of the research project, to cover a wide range of topics which facilitate the elicitation of the data. Also, the planned questions are arranged so as to make the conversation more natural as it progresses, to minimise the effect of

[63] Labov defines the observer's paradox in the following terms: 'the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation.'

the observer's paradox, which suggests that the speakers tend to modify their speech when they are monitored towards a more careful style (Labov, 1972: 113).

As mentioned earlier, the primary goal of the sociolinguistic interview is to achieve an approximation to the vernacular,[64] thus the interview considers two main conversational modules (Labov, 1984: 33-34): (1) how they see and live the process of acquiring English as a foreign language, and (2) general personal experiences; the first comprises questions related to their experiences and opinion as students of English and the second relates to their childhood, likes and dislikes, and their traditions and holidays; this second part attempts to obtain the most casual speech, as indicated by Labov (1984). The first module was included even though the elicitation of their experience of English in a formal setting (as students) might have a negative impact on the formality of the conversation, that is, the speakers might see the interviewer as a reminder of their acquisitional process, with all its implications (especially oral assessments, which are particularly stressful for students); on the other hand, this topic is the most familiar for every speaker, so their answers could be expected to be more extensive. Also, it is supposed that this conversational module is common ground for both the interviewer and the interviewee, as both parties shared the experience of being educated in the same programme, so the shared instructional experience might facilitate a less formal interaction, as the speakers might see the interviewer as a peer more than as in a hierarchical relationship (for more information on the interviewer refer to Section 3.3.3.3 below on the role of the interlocutor; for more on the hierarchical relationship between the actors of the interview cf. Labov 1984: 40-41); the second module is the one that traditionally has been used to elicit the most colloquial style from speakers – therefore, the style in which more variable forms take place[65] – (in the literature, this module has usually incorporated the "danger of death" or similar questions; cf. Labov, 1984; Milroy and Gordon, 2003; Tagliamonte, 2006: 43-44). For fifth year students, a battery of questions related to their experience as teacher trainees and how they see the teaching of English in Chile was included.

[64] Given the non-native nature of the speakers in this particular context (EFL), the vernacular is not the objective, but an approximation to informal speech.

[65] The design of my research does not consider the factor "style", as I believe this could be overridden by the fact that neither the interviewer (me) nor the interviewees were native speakers of English.

The fact that the questionnaires are pre-designed allows the researcher to modify in a certain way the questions asked, depending on the answers the informant gives; yes or no questions are avoided, but if the speaker answers briefly he/she is asked to add more information to their answer (e.g., using phrases such as *How come? Why do you say so?*, etc.). On the other hand, the use of the questionnaires has been mentioned as potentially problematic in the course of an interview, as it obstructs the natural flow of a normal interaction, in which questions arise from the information provided by the speakers (Wolfson, 1976: 192).

The questions were motivated by two of the goals of a sociolinguistic interview described by Labov (1984: 32): the obtainment of "comparative responses to questions that define contrasting attitudes and experiences among various sub-cultures", in the case of the first module the attitude towards English as a language and their personal experiences during the process of acquisition of English, and second, the elicitation of personal experiences outside their instructional setting; this second goal will serve the purpose of minimizing the observer's paradox (Labov 1970: 32). In a way, this contradicts what is expected from the speakers of the sample, that is, as teacher trainees, as they are expected to produce a "correct" Standard English, first for educational purposes (they will eventually teach English to high school students) and second, the English they acquire is based on RP pronunciation (see Section 3.2.1.1.2 on accent of instruction), which they are asked to replicate, particularly in oral examinations (e.g. they have to read phonetic transcriptions and they have to attend an interview with a board of lecturers as part of their term assessment requirements). The implications of this might be, hypothetically, the use of more standard, formal features in opposition to less vernacular, informal features (at different linguistic levels: phonetic, syntactic, discursive, etc.). To this it must be added that students have very few opportunities to speak English in real life situations, which turns out to be highly relevant in the acquisition of informal discourse features; as mentioned earlier, the fact that the interviewer (myself) is not a native speaker of English also might have implications in relation to this: first, the conversational situation is not totally natural, which might replicate the instructional circumstances (I acquired the language in the same context as the students), potentially affecting the level of formality of the language used (for more on the role of the interlocutor, see Section 3.3.3.3 below). Hence the importance of including a second conversational module, such as the one defined above, since it is

in questions such as (7) and (8) that speakers tend to use language in a more casual style (Labov, 1984).

(7) Do you have any special childhood memory?

(8) How do you usually spend holidays (national holidays, Christmas, New Year's Eve)? Which one is your favourite?

Having defined the conversational modules and the questions for each of them, I now turn to the actual interview and its practicalities, such as the equipment used, the place where the interviews were conducted, and the role of the interviewer in the development of the conversation.

3.3.3 Interview practicalities

3.3.3.1 Equipment

Many researchers have noted that the use of equipment is potentially problematic, as it reminds the informants that they are taking part in a recorded interview, making the use of natural speech even more complex (Labov, 1972; Tagliamonte, 2006). For this reason, the selection of appropriate equipment is essential to reduce the anxiety of the interviewee and, thus, achieve a more relaxed conversation.

The interviews were recorded using a portable MiniDisc Recorder Walkman MZ-R700 and a lapel microphone. The use of a minidisc recorder was preferred to any analogue counterpart, as it generally provides good quality recordings that could be easily handled, identified and digitalised; this technology, now outdated, was the best available alternative at hand at the time of the interviews. Similarly, the use of lapel microphones is much more inconspicuous than, e.g., studio microphones, as bigger microphones might intimidate speakers when placed in front of them, as it reminds them that they are being recorded. Also, as lapel microphones are placed nearer to the mouth of speakers, they pick up the voice of the speaker much more clearly, especially in high-noise environments; this last characteristic is particularly important when analysing phonetic variables, as quality is determinant for the transcription process and for the correct identification of the sounds under study.

3.3.3.2 Place of the interview

The place in which the interviews were recorded is a crucial element to be considered, since it is imperative to give the interviewees a comfortable and relaxed environment in order to reduce their anxiety about the equipment and the interview itself.[66] Wolfson (1976: 189) suggests that interviews, whether formal or informal, are pre-determined speech events, occurring in a hierarchical relationship with one interviewer asking questions and an interviewee answering them, with no possibility to choose the topic (especially in the case of interviews with pre-designed questionnaires, as in my own). Labov (1966, 1984, 2001), despite determining that the aim of the sociolinguistic interview is to elicit vernacular speech (as it is the style in which variation is more evident and easily found), also contributes to the assumption that interviews are, in fact, the observation of a speaker's linguistic performance (the objective of a sociolinguistic interview), and, therefore, there will always be a degree of care in the use of the language by the speaker, that is, the speakers will still be more careful or conscious about their speech in a language observation situation (Schilling-Estes, 2007; see also Labov, 1984; Milroy and Gordon, 2003). In the case of this group of speakers, the anxiety might also be motivated by their NNS status, that is, a natural conversation for them would take place in Spanish and not in English, and as learners of English, interviews might remind them of the oral examinations to which they are constantly subjected.

The different venues for the interviews were chosen by the interviewees themselves: their home, my home, campus (my office, a study room, and the phonetics laboratory), and the garden of a hospital (four interviews were carried out in Santiago, and in two of them the speakers and I agreed to meet at a central location; anecdotally, they chose to meet at a hospital). There are several implications in relation to the setting of the interviews, that is, it is very different to meet at a place suggested by the interviewee (their home, for instance) than meeting in a place where they do not feel very comfortable. In the last case, the implications of meeting at university might be the same as mentioned above in relation to the conversational module on English: the fact that they are talking about their process of acquisition or experiences as NNS of English in the exact same setting where their instruction took

[66] For instance, Tagliamonte (2006: 45) suggests that the ideal place for an interview (in a house) is the living room; also, she suggests that the place should avoid the presence, or the facilitation of, noise.

place might elicit a less relaxed environment and a more formal use of the language. Another implication of meeting in public places is that there is always the chance of being interrupted; interruptions in the conversation also serve the purpose of eliciting less careful speech (Chambers, 2003: 6), but in the case of the interviews for my research, these were mostly conducted in Spanish, causing the speakers to code-switch, that is, to change from English to Spanish, and then back to English, distracting them instead of facilitating the use of a more casual speech. Thus, given that circumstances prevented every interview from being conducted in an ideal place, the role of the interlocutor is fundamental to obtain the best of the speakers' performance.

3.3.3.3 Role of the interlocutor

The aim of the sociolinguistic interviewer is to obtain, from the interviewee, the largest amount of data to analyse (Labov, 1984, 1972; Schilling-Estes, 2007). Schilling-Estes (2007: 173) suggests that, despite the fact that, ideally, the interviewee is to lead the conversation, it is the interviewer who actually takes control of the interview, as "they ask most of the questions, decide which questions are asked and in what order, and are usually responsible for initiating and terminating the interaction" (*ibid*.); this means that the actual role of the interlocutor is to guide the conversation in order to obtain the most from the speaker's natural language performance. In order to achieve this, the literature describes two similar phenomena: the first is known as the *interviewer accommodation* (Trudgill 1986), in which the interviewer accommodates their speech to the interviewee; it has also been suggested that the speaker modifies their style of speaking according to the people they are interacting with or other circumstances from the environment (Milroy and Gordon, 2003; Schilling-Estes, 2001, 2007). The second phenomenon, in particular, could be considered potentially problematic, as if the interviewer modifies their speech to fit the interviewee's, his/her language performance would not be as natural as with, e.g., a pair (in case there is a hierarchical relationship). Bell and Johnson (1997) studied the role of interviewer during the interaction by switching the interviewer, matching them with the interviewee in ethnicity and gender. The results confirm that the interviewees, in fact, accommodate their speech to their interlocutor.

Another implication of the role of the interlocutor in the interaction, in the case of my study, is the fact that I am not a native speaker of English, which, in turn, creates a non-natural setting for the interview.[67] For that reason it would have been ideal to use a native speaker of English to facilitate a more natural conversation and assess whether the speakers accommodate to their speech or not, but difficulties to find native speakers willing, and being able to be part of all recording sessions did not allow it.

As lecturer and former student in the university were the sample was taken, most students knew me in advance, particularly those from third and fifth years. In the latter group, some students were contemporaneous with me (Andres, Carmen, Miguel and Carola),[68] and I supervised one of the student's undergraduate thesis (Andres); as opposed to the fact mentioned in the previous paragraph (i.e., the non-natural setting in which the students, NS of Spanish speak to another NS of Spanish in English),[69] our previous interaction as acquaintances would facilitate the interview towards more natural speech, minimising the effect of the observer's paradox. With first and some third year students, I had to introduce myself as part of the staff so that they could trust that the project had an academic purpose – that is, that the data gathered was to be used in serious academic research and not in any other situation. Furthermore, to obtain less formal speech, I was sure to mention that I was a former student (and, at the time of the interview, no longer employed as a lecturer at the university), that the current staff of Universidad del Bío-Bío would not have any access to the recordings, and that their English would not be under scrutiny. As the purpose of the study was the examination of English as a foreign language, an emphasis was made on keeping the conversation in English as relaxed as possible.

3.3.4 The interviews

The eighteen interviews took in total 12 hours and 54 minutes, and lasted an average of 41.8 minutes – the shortest being 21 minutes and 17 seconds, and the longest 77

[67] A natural setting for two native speakers of Spanish would be that they are interacting in Spanish; as they are learners of English, it would not be totally unnatural to speak to a native speaker of English, as it is part of their acquisitional process.

[68] We did not share any classes, but I was present in the department as a senior student and later as a lecturer.

[69] This group of students is used to interacting in English with NNSs of English. See Section 3.2.1.1.1.2 for greater details on the teaching of English in UBB.

minutes and 12 seconds. From these, a corpus of approximately 87,000 words is obtained.

The interviews were carried out entirely in English; Spanish was used only in the collection of personal data (not tape recorded) and code switching rarely occurred, most often when a particular lexical item was required (e.g., *"¿Cómo se dice…?*, *How do you say…?*) or when these items were not translated into English.[70] Taking into consideration the guidelines proposed by Labov (1984), Tagliamonte (2006) and Schilling-Estes (2007) in relation to the role of the interlocutor and other aspects of the sociolinguistic interview, the conversations moved around the questionnaires and speakers provided valuable information about different aspects that were later incorporated, e.g., in the indexation of the socioeconomic groups (see Section 3.4.3.2.3). As mentioned earlier, they also provided valuable opinions about their preferred accent, their beliefs on the possible effect of media on their instructional process, or their personal expectations about the whole process of acquiring English as a second language. All these, as said before, could not be quantifiable in the terms in which this research was defined. The interviews were later transcribed by me, with almost no passages marked as incomprehensible, and following a strict transcription protocol. Issues related to the transcription process and the consequent coding are presented in the following section.

3.4 Manipulation of the data

3.4.1 Introduction

The methodological procedures that follow the sociolinguistic interview have been defined and described by many authors, such as e.g., Labov (1972, 1984), Chambers (2003), and Tagliamonte (2006). The following section describes these procedures from the recording of the interviews to their analysis.

[70] These correspond mostly to cultural items, names of places, and food.

3.4.2 From transcription to coding

3.4.2.1 Transcription

Once the recordings were completed, the interviews were digitised and later transcribed orthographically; despite the fact that the variation of this study is at the phonetic level, the decision to make an orthographic transcription of the data comes from the necessity to facilitate the analysis and the extraction of the tokens, as it is considered the simplest symbolic representation of the speech file (Goedertier *et al.*, 2000); by doing this, the transcriptions are not only reduced in terms of volume, as it avoids multiple entries for each variant realisation of each of the selected features,[71] which in turn may obscure the data, but also, and as a consequence of the latter, facilitates the "automated treatment of the data" (Poplack, 1989: 431, see also, Preston, 1985; Poplack, 1993; Tagliamonte, 2006). In other words, orthographic transcription is the simplest and most practical form of speech representation that can be used for the analysis of different language levels.

The transcriptions were carried out using CLAN (Computerised Language ANalysis, MacWhinney, 2000, 2013) software designed and used to transcribe and edit speech, which allows the researcher to link the written records to audio (and video) files, facilitating the search for specific instances of the chosen variables (for a complete account of this software definition and manual, go to: http://childes.psy.cmu.edu/manuals/clan.pdf).

As this is an orthographic transcription, no interpretation of the speech was made, that is, if the speakers made mistakes in the use of grammatical forms or they used incomplete words or sentences, these were transcribed as used by the speakers. If annotations were required to help the researcher understand the speech, these were correspondingly marked as established by convention (e.g., using parentheses).

The most used transcription conventions taken in the transcription process are described next.

[71] See Section 3.4.2.1.1 for details about the transcription process.

3.4.2.1.1 Transcription conventions: the transcription protocol

The transcription, coding, and analysis require absolute rigour, and these procedures are the key steps to complete a trustworthy data set. Therefore, it is essential to establish practices that facilitate the reading and access to a well-elaborated corpus. Here, I introduce the transcription conventions utilised.

Following the habitual variationist convention of rigorously recording every single step of the research (see e.g., Tagliamonte, 2006), I designed a transcription protocol to ensure systematicity and standardisation of notations in the transcription process.[72] A master transcription was done first; this is a verbatim record of the speakers' performance, without including yet any form of phonetic variation. Once the potential variables were defined, these were sought in the transcription and marked correspondingly (see Section 3.4.2.2 for the identification and coding of the variants and variables). The protocol includes:

- *Pseudonym assignment*: to ensure the confidentiality of the speaker, a pseudonym was used in the analysis of the data and in the presentation of the results.
- *Interview register number*: each interview was tape recorded and assigned an interview register number made up of: sex (M/F) – years of instruction in English (1, 3, 5) – number of interview within the category "years of instruction in English" (1-6) – number of interview (01-18) – three first letters of the speaker's real name (e.g., M1n1-02 FEL).
- *False starts*: the occurrence of false starts is registered for both full and half words with two hyphens (*thi--, this*--).
- *Hesitation marks*: four forms of hesitation marks were identified [hm], [um], [e^{r}] and [ah].
- *Pauses*: pauses were identified with an ellipsis *(...).* If the pause was too prolonged, it was marked as *(long pause).* Long pauses were considered any pause that extended over 5 seconds.

[72] Labov provides a complete guideline for the transcription process in http://www.ling.upenn.edu/~wlabov/L560/Transcription_guidelines_FAAV.pdf.

- *Contractions*: contracted forms were marked with the stem form, the apostrophe and the contracted form: *they're, we're,* etc. The contracted forms correspond to Standard English notations.[73]
- *Non-standard uses*: the use of non-words such as *stromach* (when referring to the word *stomach*) was written as such and the intended or alleged word was written in parenthesis. In the case of the individual variable files (that is, following the transcription of the master transcription copy), the use of non-standard phonetic forms was marked in parenthesis, with the corresponding code assigned; that is, for instance for the variable voiceless alveolar fricative (ʃ), one of the variants was the voiceless alveolar affricate [ʧ] to which I assigned the code "c", thus, when the word *English* was pronounced ['ɪŋ.glɪʧ] the code "c" was written in parenthesis next to the full word: *English (c)*

A sample transcription except of the master copy (that is, with no marking of the variables examined) is presented next:

> *Interviewer: How different is, was, university from high school? About your responsibilities, your teachers, your classmates...
>
> *Carola: oh, let's see, I, I think that I consider, let's say, personally, I think that I like my time in high school more than I like, um the time in the university, mainly because my best friends now are the friends I got, I mee- I met in high school, so I have three friends and they are my best friends, I have good teachers, I had good teacher, I, I have good teachers at university also but that I think that is diffic-- different, the relationship you have when, a, with a teacher from high school is completely different, you, the one you have with a, a teacher in the university, so, and personally I think, um with my relationships I like my time in high school more than the university, but if I speak about the things I learn, the, the new people I met, um, the teachers at university I think they are quite good, at that time the university was a time for to learn, for learning, a time when I knew lots of things I didn't knew, I didn't yes, and lots of th- yes, this is I think this is the difference I can make, school was school for meeting people, for having fun, for going to parties, for going I don't know, somewhere, to the countryside

[73] When the data was preliminarily examined to discover other possible lines of research (e.g., morpho-syntactic variation) it was discovered that the speakers of the sample were highly standard in the use of morpho-syntactic features, with very little variation.

> with my friends, something like that, university was time for learning, and for, and for learning, yes, there are lots of people different for you, because I think that university you can meet so different people, you can meet people from different, er, um, from another religions, people from different countries sometimes, people who think, who think completely different from you, sometime is hard, to, I don't know, to, to keep near from them, yeah, tolerance, you have to use all your tolerance for that, and this is, university is, is a time for learning I think.

After I completed a clean transcription of each interview, I started to look for the variables to analyse. The result was the identification of four different variables, each with a different number of variants; the coding of these and of the conditioning factors is described next.

3.4.2.2 Coding

The searching of the variables was made on the basis of the function of the research questions, that is, first to test whether the variation shown by Chilean speakers of EFL is systematic (and if this is attributable to social and/or linguistic factors), and second, if the variation is similar (or different) to that shown by NS of English in relation to the same variables. The variables were not defined a-priori, as I was not sure whether there would be enough data to conduct the analysis; their selection was primarily motivated by my knowledge (and intuition) as a native speaker of the L1. In other words, I searched through the corpus for evidence of apparent variation and then I looked for the variants.

A new copy of the file was made for each variable and in each of them I exhaustively searched for and coded every variant, utilising a concordance programme (AntConc; Anthony, 2005), which allows the researcher to identify easily the feature (the variant in this case), its number of instances and the immediate context in which it appears, amongst other functions. This software was used twice: the first time to look for the orthographic representation of the sounds of interest as in the example (9) below[74]

[74] For this particular example, one of the orthographic representations of voiceless alveolar affricate is "ch", therefore, I looked for every "ch" in the corpus; exceptions (such as *Chile, chef*, and the like) were discarded. For further detail on this sort of procedure, see Chapter 4, Sections 4.1.2.5, 4.2.2.4, 4.3.2.5, and 4.4.2.5).

(see Chapter 4 and Appendix E for more on the code assigned to each variant); and the second time for extracting every occurrence of the selected feature (as shown in Figure 3.1 below); both kinds of search are made automatically by using the search tool provided by the software.

(9) a. Melinda: mm. I don't know. You say it's a special *childhood* [w] memory.
b. Josefina: *childhood* [x] memory, like a good memory or bad memory?

The following figure (3.1) shows an example of the search of one variant [x] for the variable voiceless alveolar affricate /tʃ/; the corpus files analysed (on the left frame) correspond to all female speakers from the sample; each line in the main window shows an instance of the variable (in blue, in square brackets) and the context associated with a speaker (and the speaker file, on the right frame). This process is repeated for every variant, as determined by the principle of accountability, in which "all cases where the variable element occurs in the relevant environments" are considered (Labov, 1972: 72).

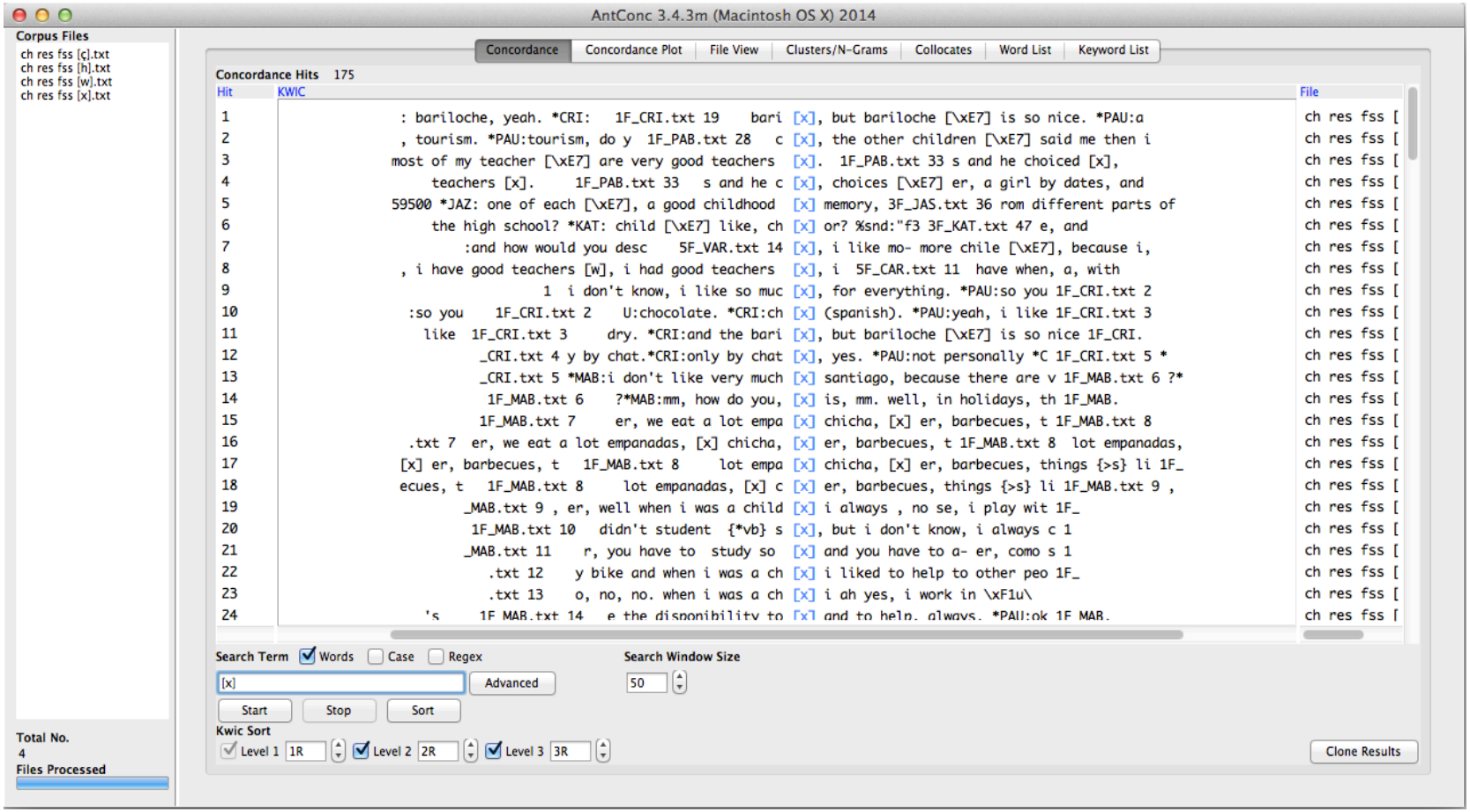


Figure 3.1 Example of the extraction of the variant [x] for voiceless alveolar affricate, for female speakers.

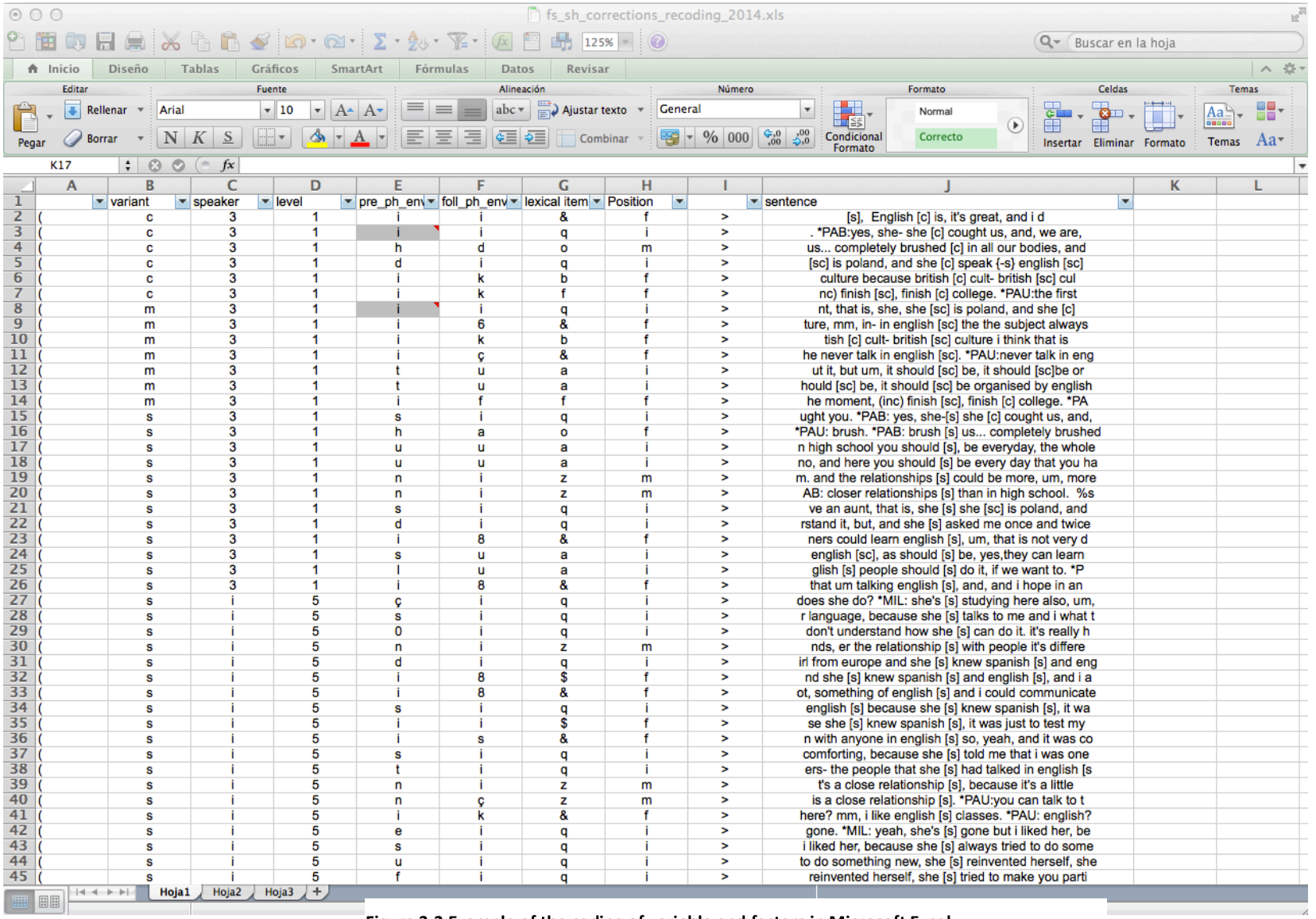
After the extraction of the variables, a coding system was elaborated to associate each of the internal and external factors to the variable and, thus, facilitate the analysis (Appendix E contains details for each code assigned to each factor group and variant, with examples). Each code assigned was entered into software like Microsoft Excel; here, each cell contains a different code for every instance and factor considered, as in the Figure 3.2 on the next page.

In this example, the context in which the variable occurs is entered, identifying the variants along with the factors considered in the design of the research (in the example, the factors are: years of instruction in English or level, preceding and following phonetic environment, lexical item, language and position); also the speakers are identified. The result from this is a string like in (10) below, which contains codes for all the factors plus the sentence in which the instance is recorded.

(10) (çr1a6mEe well, i like so much [ç] the English, because is (…)

Thus, following the coding instructions for the variable (ʧ) (see Appendix E – the codes in this particular example are given in parentheses, following the order in which they appear in the string), we learn that *much* is produced as a voiceless alveolar affricate [ʧ] (ç), by the speaker Cecilia (r), who is a first year (1). The preceding phonetic environment for [ç] is an open front vowel /a/ (a), and the following phonetic environment is the voiced dental fricative /ð/ (6).[75] The word in which the variant occurs is *much* (m), it is an English word (E), occurring at the end of the word (e).

[75] The preceding and following phonetic environments are determined from the acoustic analysis of the variable.



	A	B	C	D	E	F	G	H	I	J	K	L
1		variant	speaker	level	pre_ph_env	foll_ph_env	lexical item	Position		sentence		
2	(	c	3	1	i	i	&	f	>	[s], English [c] is, it's great, and i d		
3	(	c	3	1	i	i	q	i	>	. *PAB:yes, she- she [c] cought us, and, we are,		
4	(	c	3	1	h	d	o	m	>	us... completely brushed [c] in all our bodies, and		
5	(	c	3	1	d	i	q	i	>	[sc] is poland, and she [c] speak {-s} english [sc]		
6	(	c	3	1	i	k	b	f	>	culture because british [c] cult- british [sc] cul		
7	(	c	3	1	i	k	f	f	>	nc) finish [sc], finish [c] college. *PAU:the first		
8	(	m	3	1	i	i	q	i	>	nt, that is, she, she [sc] is poland, and she [c]		
9	(	m	3	1	i	6	&	f	>	ture, mm, in- in english [sc] the the subject always		
10	(	m	3	1	i	k	b	f	>	tish [c] cult- british [sc] culture i think that is		
11	(	m	3	1	i	ç	&	f	>	he never talk in english [sc]. *PAU:never talk in eng		
12	(	m	3	1	t	u	a	i	>	ut it, but um, it should [sc] be, it should [sc]be or		
13	(	m	3	1	t	u	a	i	>	hould [sc] be, it should [sc] be organised by english		
14	(	m	3	1	i	f	f	f	>	he moment, (inc) finish [sc], finish [c] college. *PA		
15	(	s	3	1	s	i	q	i	>	ught you. *PAB: yes, she-[s] she [c] cought us, and,		
16	(	s	3	1	h	a	o	f	>	*PAU: brush. *PAB: brush [s] us... completely brushed		
17	(	s	3	1	u	u	a	i	>	n high school you should [s], be everyday, the whole		
18	(	s	3	1	u	u	a	i	>	no, and here you should [s] be every day that you ha		
19	(	s	3	1	n	i	z	m	>	m. and the relationships [s] could be more, um, more		
20	(	s	3	1	n	i	z	m	>	AB: closer relationships [s] than in high school. %s		
21	(	s	3	1	s	i	q	i	>	ve an aunt, that is, she [s] she [sc] is poland, and		
22	(	s	3	1	d	i	q	i	>	rstand it, but, and she [s] asked me once and twice		
23	(	s	3	1	i	8	&	f	>	ners could learn english [s], um, that is not very d		
24	(	s	3	1	s	u	a	i	>	english [sc], as should [s] be, yes,they can learn		
25	(	s	3	1	l	u	a	i	>	glish [s] people should [s] do it, if we want to. *P		
26	(	s	3	1	i	8	&	f	>	that um talking english [s], and, and i hope in an		
27	(	s	i	5	ç	i	q	i	>	does she do? *MIL: she's [s] studying here also, um,		
28	(	s	i	5	s	i	q	i	>	r language, because she [s] talks to me and i what t		
29	(	s	i	5	0	i	q	i	>	don't understand how she [s] can do it. it's really h		
30	(	s	i	5	n	i	z	m	>	nds, er the relationship [s] with people it's differe		
31	(	s	i	5	d	i	q	i	>	irl from europe and she [s] knew spanish [s] and eng		
32	(	s	i	5	i	8	$	f	>	nd she [s] knew spanish [s] and english [s], and i a		
33	(	s	i	5	i	8	&	f	>	ot, something of english [s] and i could communicate		
34	(	s	i	5	s	i	q	i	>	english [s] because she [s] knew spanish [s], it wa		
35	(	s	i	5	i	i	$	f	>	se she [s] knew spanish [s], it was just to test my		
36	(	s	i	5	i	s	&	f	>	n with anyone in english [s] so, yeah, and it was co		
37	(	s	i	5	s	i	q	i	>	comforting, because she [s] told me that i was one		
38	(	s	i	5	t	i	q	i	>	ers- the people that she [s] had talked in english [s		
39	(	s	i	5	n	i	z	m	>	t's a close relationship [s], because it's a little		
40	(	s	i	5	n	ç	z	m	>	is a close relationship [s]. *PAU:you can talk to t		
41	(	s	i	5	i	k	&	f	>	here? mm, i like english [s] classes. *PAU: english?		
42	(	s	i	5	e	i	q	i	>	gone. *MIL: yeah, she's [s] gone but i liked her, be		
43	(	s	i	5	s	i	q	i	>	i liked her, because she [s] always tried to do some		
44	(	s	i	5	u	i	q	i	>	to do something new, she [s] reinvented herself, she		
45	(	s	i	5	f	i	q	i	>	reinvented herself, she [s] tried to make you parti		

Figure 3.2 Example of the coding of variable and factors in Microsoft Excel

The decision to code for the specific set of conditioning factors – which are presented in Section 3.4.3.2 below – is mainly motivated by the results obtained in previous studies, thus, e.g., in the case of voiceless alveolar fricative, previous studies of Spanish speakers (speaking Spanish) suggest that this feature is conditioned by social class and gender; but what if these factors cannot be correlated to the variation exhibited by the use of the variable in English? Thus, I standardised the coding system for the four variables, considering the same factors for all of them; that is, I include the factors already found in the literature that are specific to the selected variables, plus those that, despite not being found for a particular variable, are included in other variables. This process might also serve the purpose of finding a general pattern amongst the four variables, that is, it looks to test if the phonetic context in which the variable occurs, for instance, has the same constraining effect on the four features under study.

3.4.3 Circumscribing the variable context

The aim of this section is, essentially, to present those contexts in which the variable occurs and the elements of the analysis that are to be included and excluded (Labov, 1972; Tagliamonte, 2006). Thus, this section comprises all elements that are not considered in the analysis (*exclusions*) and all factors that are analysed.

3.4.3.1 Exclusions

Given the design of this study, which considers four different phonetic variables, the exclusion of most elements is related directly to each of the four features. However, some general guidelines are applied; the analysis excludes as context:

- Lexical items from the L1, as the focus of the study is on the use of English and not Spanish.
- Borrowings (i.e., lexical items taken from another language) and exceptional uses on the pronunciation of certain words in relation to their orthographic representation (e.g., *chef, Chicago, Chile*[76] for voiceless alveolar affricate).

[76] English is not a phonetic language like Spanish, and these contexts for the variable could be included not in the variable voiceless alveolar affricate, but in voiceless alveolar fricative; I, therefore, decided to exclude them for practical reasons, as they are considered exceptions that may obscure the results.

- Incomplete words and false starts, as, on many occasions, these contexts are re-stated by the speaker.
- Categorical contexts, in which the feature studied occurs 95%, or under 5%, of the time (Guy, 1988; Tagliamonte, 2006), as it shows that it is invariable, therefore, its inclusion in a variationist study does not make sense.[77]

The exclusions for each variable are detailed in Chapter 4. I now turn to describe all factors that are considered in the analysis.

3.4.3.2 Factor groups

Based on several studies on the four variables, the contributing factors selected in the present research can be grouped into social and internal factors, as shown in Table 3.4 at the end of this section.

All factors are considered for all variables, if viable; for instance, postvocalic /r/ is not analysed in terms of position in turn, as this variable is not found in the initial position of a turn; this particular variable is analysed, however, in relation to all other internal factors (phonetic environment, word frequency, and word class).

3.4.3.2.1 Speakers' sex

A recurrent factor in variationist research, gender is considered "a powerful differentiating factor in almost every case of stable social stratification and change in progress that has been studied" (Labov, 2001: 262; see also Cheshire, 2002); that is, sex has been traditionally linked in variationist research to the actual language performance of speakers in both stable and changing language situations. Despite my study being out of the scope of these language situations, I examine this factor given its relevance in the study of variation in general, and also, as it has been linked to the varying use of diverse features in both L1 and L2 situations. For instance, in L1 sex has been widely studied as one of the main factors involved in variation (refer, e.g., to Holmes, 1997; Milroy *et al*., 1999; Horvath, 1985, Milroy and Gordon, 2003, amongst many others). Furthermore, Eckert (1988) has suggested that, given the role of

[77] This context of exception is reviewed further in the Chapter 4 (e.g. Section 4.2.3.2.1.2, *Categorical speakers*).

females in society (typically patriarchal), women may have a social motivation to utilise the resources provided by language, such as their higher use of prestigious features. Similar patterning has been found in L2 research, where females have been found to favour more prestigious forms than males (Hiang and Gupta, 1992). However, Major (2004: 171) suggests that the number of variation in SLA studies addressing sex is small, but that it is commonly believed that females present higher use of target-like and prestigious forms (see e.g., Weiss, 1970). This means that sex is undoubtedly one of the most interesting factor groups to be observed.

3.4.3.2.2 Years of instruction in English (YIE)

It is suggested that the process of acquisition of an L2 is gradual and that the more the learner is exposed to the target language the closer he gets to the target forms (Flege and Hillerbrand, 1986; Flege, 1995; Flege *et al.*, 1997, Slavoff and Johnson, 1995). Most of the literature that refers to the amount of exposure to English is basically descriptive, i.e., it characterises the actual performance of a speaker (or group of speakers) in relation to a particular feature, and, therefore, the focus is on the description of the acoustical characteristics of learners' performance and not on determining patterns or sources of variation.

Furthermore, Flege *et al.* (1997: 438) suggest that in the case of acquisition of English by Spanish speakers it is "related inversely to their age of first extensive exposure to native-produced English". In the case of my research, the exposure of the speakers to English cannot be considered extensively in the terms described by the literature (often referring to immigrant settings), but instead it refers to the amount of time of formal instruction in English, that is, their interaction with the language is cumulative in years of experience in the instructional setting, with very limited access to native speakers from which they could acquire native-like forms from real life interaction situations (see Section 3.2.1.2.2 for more on the description of this group of speakers' acquisitional process).

The literature suggests that in cases where the individuals were exposed to formal instruction and then were immersed in the L2 culture, they acquired more native-like features than those who were not exposed to the immersion process (Regan, 1996); also in immigrant settings, their use of native-like features increases according to the

years they have been in the L2 setting (Flege, 1995), but how will this be the case with those speakers who are not exposed to English in a natural setting, as the speakers of this sample? As we cannot expect the occurrence of native-like variation in non-naturalistic settings, we might count on that along with a greater exposure to the L2, these learners will be able to produce, progressively, more target-like features.

3.4.3.2.3 Socioeconomic group: definition and indexation

3.4.3.2.3.1 Introduction

Even though social class was not considered a requirement to define the sample, it is included as a potential conditioning factor, as traditionally in variationist research it has been noted as one of the most important social factors perhaps that contribute to linguistic variation and it implies a subscription to certain norms of behaviour imposed by our environment, which include language use, i.e., the way we talk seems to be a clear indicator of our social background; for instance, many studies in L1 situations have shown that the lower social classes tend to use higher rates of non-standard forms, whereas the highest rates of standard uses are found in the upper classes (see e.g., Trudgill, 1974; Guy, 1989; Labov, 2001; Romaine, 2008). An interesting phenomenon commonly found in relation to the variation exhibited by middle class speakers is *hypercorrection,* in which members of this social class tend to show higher rates of prestigious features, even higher than upper class speakers (see, e.g., Decamp, 1972; Labov, 1966, 1972, 1990, 2001; for more on hypercorrection, see Chapter 4, Section 4.1). In L2 situations, the effect of social class has not been as widely studied as, e.g., style or gender; however, results are similar to NSs situations, that is, a social class differentiation in the use of prestigious and non-prestigious features has been found, with an interesting effect on the status of the feature under study in the L1 and the L2 (Beebe, 2006). Thus, the effect of social class on the varying features calls for its inclusion in the analysis, given the tested effect of class in L1 and L2 situations. The next section describes how the social class of the speakers of my sample was indexed.

3.4.3.2.3.2 Indexation

As mentioned before, the selection of the speakers of the sample was not based on the social class of the speakers, that is, I did not look specifically for a certain number of speakers from each identified socioeconomic group, but the adscription of the speakers to social class was made a posteriori, based on the information given by the participants of the sample.

The elaboration of this index was based on a classification of the Chilean socioeconomic groups as determined by a Chilean marketing research company (ADIMARK),[78] usually used as a reference for official data. Labov (2001: 60) suggests that "societies are stratified by differential access to socially valued resources, which can be measured objectively"; this measurement gives as a result the social categorisation of individuals into social classes, based on factors such as "income, house type and dwelling area" (Labov, 2001: 60, reporting on the *Index of Status Characteristics* elaborated by Wagner, 1960), education, occupation and residence value (Labov, 1980, 2001, reporting on the project *Linguistic Change and Variation*).

Unlike in other studies, I do not elaborate on a socioeconomic index assigning specific values to every category included in the index, but I took as a reference the current social classification and, based on the information provided, assigned the informants to one of the five groups. A similar procedure was followed by Sadowsky and Salamanca (2011); Sadowsky (personal communication) adapted the model described below (five economic groups from A to E), incorporating both parents' educational background and occupation (not only the breadwinner, as he/she might not be the greatest linguistic influence on the informant).

The groups are classified and defined as follow:

- ABC1 (upper class): divided into a set AB (upper-upper class, 2.5%) and C1 (upper class, 97.5%), this group represents 10% of the population. The income varies from £1,900 to £8,450 GBP monthly. Regarding occupation, this group holds mostly high executive positions, and its individuals hold at least one higher educational degree (≥18 years of instruction).

[78] ADIMARK uses the index ESOMAR, originally designed by World Association of Market Research

- C2 (upper-middle class): this group represents less than 20%. The income varies from £708 to £1,900 GBP monthly. In terms of occupation and education, the members of this group hold higher (college and university) degrees (15 years of instruction in average) and work in executive positions (young professionals, executives, small businessmen in industrial and commercial companies, accountants, teachers, technicians, middle level employees). This group usually owns at least one small house or flat, and one or two less exclusive car(s) (in relation to ABC1 individuals).
- C3 (middle-middle class): represents about 30% of the population, with incomes averaging £570 GBP monthly. Even though this group includes public and private employees, teachers, specialised technicians and workmen, artisans, small businessmen, salesmen, and insurance salespeople, most of the individuals do not hold higher educational degrees[79] (13 years of instruction in average). This group can afford small houses and flats in less privileged neighbourhoods, depending on mortgage rates, and also the rental rate in this group is considerably high. The vehicles owned in this group are usually used for working (taxis, vans, trucks, and the like).
- D (lower middle class): representing more than 30% of the population, this group earns from £258 to £464 GBP monthly. Workmen, low level employees, messengers, and cleaning employees belong to this group, who do not hold complete elementary or secondary education[80] (an average of 11 years of instruction).
- E (lower class): represents less than 7% of the population and in this group the average educational level is not higher than fifth elementary (≤ 8 years of instruction). The average income is £126 GBP monthly, from occasional, informal work.

The indexation of the social class was completed incorporating into the groups A to E the information provided by the speakers of the sample; thus, for instance, Paola completed elementary and secondary education in public schools, both her parents

[79] Up to 1974, elementary teachers were not trained in universities, but in "Escuelas Normales", i.e., educational institutions that specialized in teacher training. No university in Chile has technical degree programmes, which are taught in "Institutos de capacitación" y "Centros de Formación Técnica", the equivalent to colleges in the U.K.

[80] The Chilean compulsory educational system is divided into elementary (6-13-year-old students in 8 levels) and secondary (14-18-year-old students in four levels) education.

completed secondary education, her father works as a driver and her mother as a housekeeper; therefore, she belongs into group D; the information for the educational level and occupation of both parents for each speaker is provided in Table 3.2. It has been argued, however, that the fact all speakers are in university means that they may present a different use of the linguistic variables from the actual group they belong to. For instance, Labov (2001: 115) suggests that the educational level of the family cannot be automatically ascribed to a child that has not completed his education for obvious reasons: if it is considered that an upper class child has attended only up to e.g., first year of elementary school, his situation could be equal to an adult from a lower social class who presents with a similar educational background, making the index untrue. Labov elaborated in this way an index in which he considered separately the educational level of the parents and the educational level of the child.

In addition, it could be argued that social mobility also has a role to play in the social indexation of my sample: for definition, group D (lower-middle) does not include individuals who attend or have attended higher education; however, they are included in group D for a number of reasons: (1) the educational background of their parents, (2) the occupation of their parents, (3) their elementary and secondary educational background, (4) the fact that they have not yet completed their higher educational programme, and, therefore, they do not have an income, and (5) the kind of university they attend is also an indicator of social class. This means that, at the time of the interview, speakers were basically included in one of the social groups identified from the social conditions of their family and not on their own as individuals. Table 3.2 shows the data drawn from the interviewee data sheet that helped me identify the socioeconomic group of the speakers.

Table 3.2 Educational level and occupation of the speakers' parents		
Speaker	Father's Educational Level and Occupation	Mother's Educational Level and Occupation
Paola	SC - Driver	SC – Housekeeper
Cecilia	HTC – Mechanic	SC – Housewife
Marcela	HTC – Unemployed	HUC – supervisor (Ministry of Education)
Consuelo	SC – Accountant	SC – Housewife
Caterina	SC – Truck driver	SC – Shop assistant
Josefina	SC – no information	SC – Secretary
Melinda	HUI – agronomist	SC – Hairdresser
Carola	HTC – transportation manager	SC – Housewife
Carmen	SC – driver	SC – Housewife
Francisco	SC – delivery	SC – Housewife
Jaime	SC – mechanic	EI – Housewife
Sebastian	EC – baker	SC – Housewife
Bernardo	SC– police officer	SC – Secretary/Housewife
David	HUC– architect	HUC – Housewife
Juan	HUC– n/a (deceased)	HTC – Housewife
Cesar	SC– taxi driver	SC – Housewife
Andres	SC– farmer	SC – Housewife
Miguel	HTC – food technician	HUC – teacher

EI = elementary incomplete; EC = elementary incomplete; SI = secondary incomplete; SC = secondary complete; HTI = higher technical incomplete; HTC = higher technical complete; HUI = higher university incomplete; HUC = higher university complete

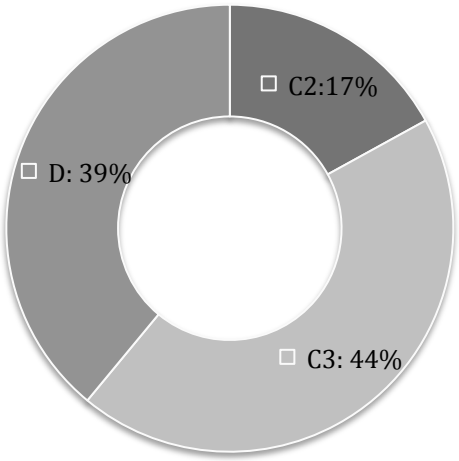
Table 3.3 shows the educational background of the speakers, another factor considered in the indexation of social class. Summing up the information given above, Table 3.3 also shows the socioeconomic group to which the informants belong.

Table 3.3 Speakers' primary and secondary educational background and socioeconomic group

Speaker	Elementary Education	Secondary Education	SEG
Paola	Public	Public	D
Cecilia	Subsidised	Subsidised	C3
Marcela	Private	Subsidised	C3
Consuelo	Subsidised	Subsidised	C3
Caterina	Subsidised	Subsidised	D
Josefina	Subsidised	Private	C3
Melinda	Subsidised	Subsidised	C3
Carola	Subsidised	Subsidised	C3
Carmen	Private	Subsidised	D
Francisco	Public	Public	D
Jaime	Public	Subsidised	D
Sebastian	Subsidised	Subsidised	D
Bernardo	Subsidised	Subsidised	C3
David	Private	Private	C2
Juan	Private	Private	C2
Cesar	Public	Subsidised	D
Andres	Private	Private	C3
Miguel	Subsidised	Subsidised	C2

Most of the speakers belong to groups C3 (N= 8, 44%) and D (N=7, 39%), as shown by Figure 3.3 below. Only three speakers belong to the group C2 (17%). In general terms, the group is pretty homogenous: all the speakers belong to one level of middle class (upper, middle or lower); further analysis on the data will reveal if the differences in speech can be attributable to social class, or any other group, or interacting factor groups.

Figure 3.3 Socioeconomic groups of speakers in the sample (N=18)



3.4.3.2.4 The role of the individual

The examination of social constraining factors would not be complete without considering the role that the individual plays on variation. Stuart-Smith and Timmins (2010) suggest that "individuals are inevitably at the heart of language variation and change, because it is the adjustments in linguistic behaviour of individuals, whether consciously or more usually subconsciously, which constitute variation, and potentially change." This means that the origin of variation is in the linguistic behaviour of individual speakers, which is subsequently extended to linguistic communities, but, as argued by Guy (1980: 2), "is variation in the speech community the result of the diversity of the group, reflecting the organisation of society into a number of discrete lects within which variation is at a minimum, or is this variation present with identical uniform structures in the speech of every individual?"; this question validates the idea of using individual speakers in the examination of the linguistic behaviour of a group, but leaves open the question of whether the variation is the mirror of the speaker, or vice versa. Guy (1980) contrasts the varying linguistic behaviour of individuals against the group, defining the methodological procedures to examine such patterns of variation; his results, however, suggest that the examination of individuals, instead of revealing patterns and regularities, may obscure them, as including individuality as a factor would multiply the amount of possible conditioning environments.

The contribution of individuality in this study is to be taken into consideration as first, the researcher must get access to the speaker and to their community (e.g., to understand the spread of variation in a language community) and second, to study how the individual relates to a major group of people, e.g., when considering aspects such as language and identity (Noels *et al.,* 1996; Holmes, 1997; Holmes and Meyerhoff, 2008), or when examining style variation or processes such as accommodation (see e.g., Milroy and Milroy, 1985; Labov, 2001; Eckert, 2005; Stuart-Smith and Timmins, 2010). However, as suggested by Labov (2001: 33-4), the study of the individual is expected to shed light on the "comparative use of resources made available by the speech community"; in other words, the individual not only contributes to the variation of the group, but he is also a mirror of the linguistic behaviour of his community of speakers.

The possible social constraints on use are reviewed for each variable in Chapter 4. I now turn to the linguistic constraints which may have an impact on the observed variation.

3.4.3.2.5 Phonetic environment

Previous research in phonetics and phonology has shown that there are many phenomena – universals and language specific – that modify the realisation of sounds, such as assibilation, lenition, palatalization, reduction, etc. (see e.g., Dickerson, 1976; Manuel, 1995; Lavoie, 2002; Local, 2003; Gordon and Ladefoged, 2001); similar processes, such as strengthening and lenition, have been found to occur in linguistic variation contexts. For instance, the lenition found by Cedergren in the Panama City speech (1973), was the result of phonetic environment and position of the variable in the word, amongst other factors.

Most variationist studies that have considered phonetic environment have shown this to be a major conditioning factor (see e.g., Labov, 1966, 1972, 1981, 1994; Fasold, 1972; Cedergren, 1973; Guy, 1980, 1991, 2007; Bybee, 2002; Raymond *et al.*, 2006; Wolfram, 1969, amongst many others). In Spanish as L1, Pérez (2007) shows that the occurrence of a given phonetic variable in intervocalic position favours the occurrence of weakening processes, a similar process found in English as L2 (Flege, 1995). The inclusion of this factor will shed light on the similarities and differences of

the effect of both preceding and following segment on the variation of NNSs in non-immersion settings; also, it will test whether this factor is as powerful in NNSs as it is for NSs.

3.4.3.2.6 Prosodic position

Two different positions are referred to in this section: within the word (initial, middle and final, e.g., for (sh) *sh*ip, a*sh*ore, bu*sh*), and in turn (specifically for (th), initial and other, as in *Th*e house is full. I saw you in *th*e house). As with phonetic environment, the fact that two language systems are interacting is highly relevant, as for all variables, there are differences in the places in which they occur for English and Spanish (see Chapter 4 for each variable in greater detail). The effect of prosodic position is recurrently linked to phonetic environment, as it has been shown, for instance, that the occurrence of particular features in coda position is strongly correlated to phenomena such as co-articulation (Huffman, 2005).

3.4.3.2.7 Individual lexical item

The effect of the individual lexical item in the study of variation has been examined mostly in relation to two lexical characteristics: class and frequency.

In relation to class, the variation of certain phonetic features has been linked to whether the words in which they occur are open or closed class words (Bell *et al.*, 1999, 2001; Raymond *et al.*, 2006).[81] Research has shown that, e.g., in NSs of English in the study of function versus content words, properties such as the stress, vowel reduction, presence of syllabic consonants, and so on, have an effect on the differentiation of content and function words.[82] The study of variation in relation to lexical items in native speakers of English has shown that, e.g., closed class words tend to vary extensively in conversational settings, being frequently shortened (from their full forms) by speakers with high proficiency, in interaction with the prosodic structure of the feature, particularly the occurrence of final lengthening and

[81] A similar word classification suggests the terms *content* and *function* words.

[82] Studies from different subfields of linguistics have studied e.g., the effect of gender on the use of function and content words as an indicator or density of information (Keune *et al*., 2005), the borrowability of function and content words (Field, 2002; Poplack, 1993), the predictability of function vs. content words (van Petten and Kutas, 1991), and brain hemisphere differences in the use of content vs. function words Kutas and Hillyard, 1983), amongst many others.

weakening and initial strengthening (Bell, 2001);[83] and that they are strengthened in the presence of pause or disruption, where the context is less predictable and depending on the position of the word in the utterance (Bell *et al.*, 2001). It has also been tested that higher rates of coronal stop deletion occur in function than in content words (Raymond *et al.*, 2006), and in interaction with frequency (Jurafsky *et al.*, 2001).

Other studies related to the frequency of works link sound changes to the high or low frequency of words (see e.g., Bybee, 2002; Santa Ana, 1991; Jurafsky *et al.*, 2001). I also study the structural similarities of lexical items, that is, similarities in terms of form, phonetic context or position of the variable in the word (see Chapter 4 for more detail on each variable).

3.4.3.2.8 Summary

The effect of both internal and external factors has been widely documented in variationist literature. The account briefly presented here is a sample of some of the most relevant studies that justify the inclusion of the aforementioned conditioning factors (a complete account of the literature for each variable is presented in the corresponding sections in Chapter 4; see e.g. Section 4.1.2 for voiceless alveolar fricative, Section 4.2.2 for voiceless alveolar affricate, Section 4.3.2 for voiced dental fricative and 4.3.2 for postvocalic (r)). Table 3.4 summarises the factors considered in the analysis, if viable.

[83] Bell *et al.* (2001) obtain the data from the Switchboard corpus of conversational speech in speakers of General American English.

Table 3.4 External and Internal factors groups considered for all variables selected.

Type of Factor	Factor Group	Factors
External	Sex	Male (M)
		Female (F)
	YIE[84]	1
		3
		5
	SEG	C2
		C3
		D
Internal	Phonetic Environment[85]	Vowels
		Plosives
		Fricatives
		Affricates
		Approximants
		Laterals
		Pause
	Prosodic Position	Turn Initial
		Other in Turn
		Word Initial
		Word Middle
		Word Final
	Lexical Item	

Having introduced the potential conditional factors for the selected variable, I now turn to one of the most relevant parts of any study of this kind: the results drawn from the analysis conducted in the four variables selected.

[84] Years of instruction in English.

[85] Both preceding and following phonetic environment.

CHAPTER 4 – RESULTS

4.1 Voiceless alveolar fricative (ʃ)

4.1.1 The variable

The voiceless alveolar fricative (ʃ) appears in a variety of linguistic contexts in English, such as *show* /ʃəʊ/,[86] *perception* /pəˈsepʃᵊn/, *commercial* /kəˈmɜːʃᵊl/ and *impression* /ɪmˈpreʃᵊn/. To the best of my knowledge, the variation of this feature has not been studied – as a variable – in native speakers of English, but it has been described in general linguistics literature (Catford, 2001; Ladefoged, 2001) and in particular, e.g., in relation to the effect of vowels on its realisation (Shadle *et al.,* 2008; Wilde, 1995). It also has been referred to in a study of (s) in Glaswegian (Stuart-Smith *et al*., 2003; Stuart-Smith, 2007), as a variant for (s) in relation to the perception of sexual identity (Levon, 2007) and in the alternation of (sh/ch)[87] in Latino speakers in New York (Labov, 2012). In Spanish, it has been shown to be one of the variants for voiceless alveolar affricate (ʧ), a variation which has been associated mainly with external factors such as gender, social class (see e.g., Cedergren, 1973, 1984; Valencia, 1993-1994). The following sections provide insights on the study of the variation of voiceless alveolar fricative and its relevance for the present research.

4.1.2 Relevant findings and hypotheses

The use of this variable has primarily been linked to social factors, with social class being one of the most salient constraints. One of the most relevant facts about this feature in my research is that, despite the small amount of studies that examine this feature in English, I selected it because of the effect this feature has in Chilean Spanish, where it is considered a stigmatised alternative of the voiceless alveolar affricate. This same alternation (/ʃ/ - /tʃ/) has a different sociolinguistic value in other varieties of Spanish (see e.g., in Panama Spanish, Cedergren 1973, 1984) and in speakers of ESL with Latino backgrounds in the U.S. (Labov, 2012), as will be detailed below. But, how can studies in Spanish be relevant in a study where the data is in

[86] This and following standard phonetic transcriptions correspond to RP.

[87] Labov (2012) uses this notation (sh/ch) for the study of the most common graphemic representations of voiceless alveolar fricative /ʃ/ (sh) and of voiceless alveolar affricate /tʃ/ (ch).

English? Recall that one of the main hypotheses for my research is the transfer of sociolinguistic patterns from the L1 (Chilean Spanish) to the L2 (English), therefore it is highly relevant to understand how this feature behaves in the L1 of the speakers of the sample. The following sections introduce previous studies on the variation of voiceless alveolar fricative (ʃ) for English, Spanish and ESL/EFL situations. At the end of the literature review that follows, I propose a series of hypotheses that are specific to this variable.

4.1.2.1 Variation in English

As mentioned in the introduction to this variable, the study of voiceless alveolar fricative in English has been mainly restricted to the description of the feature in articulatory terms, with limited approach into variationist linguistics (see e.g., Jongman, Wayland and Wong, 2000; Catford, 2001; Ladefoged, 2001; Li, Edwards and Beckman, 2009; Ladefoged and Johnson, 2011). In variationist research, it has been linked to working-class speakers from Glasgow, as an articulatory similar feature for /s/ (Stuart-Smith *et al*., 2003; Smith, 2007);[88] however, Stuart-Smith argues that "no measurement of /ʃ/ shows sociolinguistic variation" in opposition to the use of /s/ (2007: 82); this suggests that, in (Glaswegian) English the use of /ʃ/ does not have the same sociolinguistic conditioning as in Chilean Spanish. Another study in English that links /ʃ/ to /s/ is Levon's study of the perception of sexual identity, in which he uses the phoneme /ʃ/ as a control feature for the study of two prosodic variables: pitch range and sibilant duration.[89] However, these studies do not examine /ʃ/ as a variable but as a variant occurring in restricted settings; the studies show that its occurrence may be sociolinguistically linked to, e.g., social class in some varieties of English. This restricted use of /ʃ/ does not occur in Chilean Spanish; nevertheless, I examine the variation of this feature in the L1 of the speakers of the sample in the following sections.

[88] Stuart-Smith *et al.* (2003) and Stuart-Smith (2007) suggest that the articulation of /s/ in working class speakers is 'retracted', which pronunciation resembles the articulation of /ʃ/. They report on similar articulatory descriptions from Macafee (1983) and Catford (1977).

[89] In both Stuart-Smith (2007) and Levon (2007) the duration of sibilant /ʃ/ is shorter than of /s/.

4.1.2.2 Variation in Spanish

In Chilean Spanish and other varieties of Spanish, voiceless alveolar fricative does not have an orthographic representation as shown above for English. Furthermore, it is described in phonology as an allophone "in free variation and possibly [in] complementary distribution with /ʧ/" (Dalbor 1959: 67). This means that the voiceless alveolar fricative is not shown in the inventory of standard Spanish, despite the fact that its realisation is present in the intervocalic, medial position in words such as *muchacho* (standard [muʧaʧo] versus non-standard [mu'ʃaʃo], boy), after consonants as in *ancho* (standard ['anʧo] versus non-standard ['anʃo], wide) and in the initial position as in *chala* (standard ['ʧala] versus non-standard ['ʃala], sandal); in Spanish, this feature does not occur in the final position.

As a variable, (ʃ) has been documented in varieties of the Dominican Republic, Cuba, and Spain (Andalucia) (see e.g., Canfield, 1962; Lipski, 1994). Cedergren (1973, 1987) examined the lenition of (ch), which she defines as the "alternation of a fricative palatal [š] with the standard [č]"[90] (Cedergren, 1987: 51), in Panama City, in interaction with factors such as the prosodic position in the word –initial, and middle-, phonological environment, style, place of origin, sex, age, and socioeconomic class. Cedergren suggests that the alternation of (ch) is limited to Spain, the Dominican Republic, Cuba and New Mexican Spanish.

Cedergren (1973: 66) reports the use of three variants: an affricate [ʧ], a voiceless palatal affricate with a reduced stop onset [^{t}ʃ],[91], and a non-labialised voiceless palatal fricative [ʃ]. Cedergren (1973: 79) concludes that the use of voiceless alveolar fricative was a linguistic change of recent appearance, spreading country-wide from the urban area, "initiated by younger lower middle class adults"; this process is marked by the "gradual phonetic implementation along the pathway [ʧ] > [^{t}ʃ] > [ʃ]".

Cedergren revisits Panama City, analysing the features of the original data set (1982 and 1984, results presented in 1987) half a generation later, to test the status of phonetic change, which was confirmed; that is, the lenition of (ch) has spread in time and place from Panama City to other cities. The factor that seems to contribute the

[90] The Americanist phones [š] and [č] correspond to the IPA signs [ʃ] and [ʧ] respectively.

[91] My own notation, Cedergren uses $\hat{\hat{S}}$

most is age (young adult speakers), which she interprets as "an indication of their sensitivity to the social importance of the newer variant"; she compares this pattern with the hypercorrection shown by speakers in the study of New York City (Labov, 1966).

Her conclusions also suggest that the variation exhibited a consistent use of the prestigious variant based on the male/female distinction and that the reducing of occlusion (lenition) seemed to be favoured by preceding vowel and pause. Also, she uncovered the presence of a curvilinear pattern in the lenition of (ch) based on the socioeconomic group of the speakers, in which the extreme groups favoured [^{t}ʃ] and [ʃ]. Thus, Cedergren's study demonstrates that the variation of (ʃ) is the result of the interaction of social and linguistic factors, specifically of age, sex, place or origin, preceding segment and prosodic position; her work set the basis for the study of (ʃ) in Spanish, as it confirms that the use of the variable is constrained by internal and external factors in native speech.

The sociolinguistic values assigned to the variable in Panama City are, however, opposed to those in Chilean Spanish. In Chilean Spanish, the use of voiceless alveolar fricative has shown to be constrained by sociolinguistic factors, such as age and place of origin, but particularly by socioeconomic class (see e.g., Valdivieso, 1983; Tassara, 1992). For instance, Bernales (1978) notes that educational level, sex, style (casual versus emphatic) and position (initial, intervocalic, and after consonants) are all contributing factors in the realisation of (ʧ) as [ʃ]. Widgorsky (1978), Valencia (1993-1994) and Cepeda (2001) have shown that this feature is characteristic of lower-social-class subjects, whereas Valdivieso (1993) notes that male and lower-class speakers have shown higher rates of use of the non-standard variant, in this case (ʃ), than female and upper-class speakers. The linguistic and social conditioning of (ʃ) in Chilean speakers make this feature very interesting to observe, particularly in order to test whether or not they carry the social meaning of this feature from Chilean Spanish into English by using the non-target prestigious form from their L1 over the target form.

4.1.2.3 Variation in ESL/EFL contexts

The study of this feature in contexts in which English is not the first language of the speakers has revealed the occurrence of at least two variants, which match the variants exhibited by speakers of Chilean Spanish, i.e., the fricative [ʃ] and the affricate [ʧ]. In a study on a Cuban speaker studying English in the U.S., Dalbor (1954)[92] suggests that the variation of these features is the result of the difficulties the speakers experience in recognising the phonetic distinctions between them and that, with the appropriate training, the error could be corrected. Eklund and Lindström (1998) show that the inclusion of xenophones (phones from a foreign language), such as the voiceless alveolar fricative, also present variation in speakers of Swedish. Flege (1991: 256) reports that English speakers learning Spanish with native speakers of Spanish have been found to produce fricative (ʃ) as [č],[93] which suggests that English speakers are aware that the variation of this feature is characteristic of Spanish speakers. Finally, Labov (2012) reports on New York and Philadelphia speakers with four different ethnic origins: African American, Caucasian, Latino Spanish readers and Latino English readers. He studies the "alternation of (ch/sh)" in reading passages. He suggests that the occurrence of "errors" (i.e., the use of deviant forms) in the reading of words containing (ch/sh) was different for African American and Latino speakers. For the first group these errors were grammatical and for the second they were grammatical and phonological. Labov concludes that the alternation (merger) between the affricate /tʃ/ and the fricative /ʃ/ was identified as a consequence of learning to read in Spanish first then in English, but it has "no direct consequences for reading or comprehension" (Labov, 2012). This last study is highly relevant for my own, as it suggests that the L1 of the speakers has a real impact on the use of the variable under study here.

Considering the above sections in relation to the findings for English, Spanish and ESL, I formulate a series of hypotheses which are specific for the variable voiceless alveolar fricative /ʃ/; these are presented below.

---

[92] Dalbor's study was not carried out using variationist methodology, but is based on the observation of the acquisition process from an educational (EFL/ESL) perspective, in which variation is understood as erroneous performance.

[93] Cf. note 92, in this chapter.

4.1.2.4 Hypotheses

Despite this feature being one of the standard phonemes in native speaker English, (ʃ) has been considered a stereotype of lower class, uneducated speakers of Chilean Spanish (see e.g., Valdivieso, 1983; Tassara, 1992). This makes it a very interesting feature to look at in EFL contexts, as one of the hypotheses of my work is to test whether the linguistic variation exhibited by native speakers of Chilean Spanish is transferred to non-native English by means of hypercorrection. Hypercorrection may take different forms; for instance, Labov (1972: 126) defines it as a "deviation from the regularity [in the behaviour of lower class speakers] of the behaviour of other classes [generally upper]"; however, in the case of this research, the hypercorrection expected is the transfer of a form of a prestigious (not of the stigmatised) variant from native to non-native speech.

Given that the variation of the voiceless alveolar fricative is constrained by linguistic and, especially, by social factors, I propose as a general hypothesis that the systematic occurrence of variation of (ʃ) evidenced in native speakers of Chilean Spanish is replicated in non-native English by means of the transference of L1 prestigious features into English. Specifically, it is expected that the use of the English standard [ʃ] is to some extent replaced by a more prestigious variant in the L1 of the speakers, for instance the voiceless alveolar affricate [ʧ], to avoid the use of the target form, which is stigmatised in their L1, as in (11) below:

(11) Um, because I like *English* ['ɪŋ.glɪʧ] and I think *English* [/'ɪŋ.glɪtʃ/] is a good, um it's a good opportunity, for do different things

(Juan, third-year, male speaker).

Decamp (1972) extends the definition of hypercorrection to other dialects of English, other languages (Taiwanese) and cases of language contact (particularly the case of Jamaican Creole), therefore his definition is more appropriate to my study than Labov's; he considers it to be "an incorrect analogy with a form in a prestige dialect which the speaker has imperfectly mastered" (Decamp, 1982: 87), and thus, hypercorrection turns to be an inconsistent behaviour if the speaker is aware of the existence of the rule but not of its application, resulting in random simultaneous

overgeneralisation and undergeneralisation of the rule (Decamp, 1972). If hypercorrection is confirmed and if speakers are aware of the stigmatised characteristic of (ʃ), they will fluctuate between the use of the standard [ʃ] and its replacement by the Chilean prestigious form [ʧ]; given that in Chilean Spanish the use of this variable is bound to social class and education, it is expected that this particular pattern will be replicated in this group of speakers, depending on their social background; that is, there should be a different pattern in upper-middle and lower-middle-class speakers.

Previous research on this feature leads to the formulation of a number of hypotheses:

- There is a difference between the linguistic performance of male and female NNSs of English, with female speakers showing higher rates of the prestigious Chilean Spanish form [ʧ] (non-standard in English) over the English standard [ʃ] (stigmatised in Chilean Spanish). However, this particular hypothesis may have a double effect on the results, as it also has been shown that, in diverse studies across languages, females are more standard than males (see Labov's *Gender Paradox*, 2001, Chapter 8 for a full account on gendered differences in both changing and stable linguistic situations)
- The effect of the phonetic environment is a robust contributing factor to the variation, especially of vowels and pause, favouring the use of fricatives, as shown in Cedergren (1973).
- The experience in English as a foreign language is proportional to the use of target forms; that is, the more years of study, the closer the speakers are to the use of target forms. This prediction is based on the results that relate the number of years of exposure to the L2 to the closer use of target forms by NNSs (Flege, Munro and MacKay, 1995d).
- In relation to lexical items, the weakening of features in closed-class forms could be an expected result, as opposed to the strengthening of open-class forms. Thus, the use of fricatives will be higher in closed-class than open-class words and affricates (see e.g., Bell, 2001, and Bell *et al*., 2001). Also, based on previous results for e.g., phonological environment (Cedergren, 1973; Guy, 1991, 2007; Poplack *et al.,* 2000a; Poplack and Tagliamonte, 2001), words that are alike in terms of class, form, position of the variable in the word, etc.

will behave similarly (e.g., the pair *English* and *Spanish,* which share the same preceding segment and class).

- For position, weaker forms (articulatory speaking, fricatives over affricates) will show higher rates of use in middle and final position and stronger forms in initial position, as, e.g., in Bell *et al.,* (2002).

4.1.2.5 Exclusions

From the corpus, all tokens containing the variable (ʃ) were coded and extracted. No exceptions were made on the basis of the type or nature of the lexical item, such as Spanish words or borrowings, as these do not exist or were not used by speakers.[94] Nevertheless, a very large number of tokens for the lexical item *English* was noted in comparison with other lexical items[95] (507 out of a total of 924, which corresponds to 54% of the data); given the saliency of this item, and considering its amount in the data, I selected up to five tokens of *English* per speaker. This procedure has been justified previously (see e.g. Wolfram, 1969, 1993; Tagliamonte and Molfenter, 2007), and it is done to guarantee that the results are not obscured by an item which preliminary examination has shown to be rather invariable.[96]

Having introduced the specific hypotheses for (ʃ), and having circumscribed the variable context, I now turn to the results section, particularly to the introduction of the variants found in the analysis and how they are distributed. This section is followed by the contrastive analysis of my results with the literature in relation to the conditioning factors, or group of factors, selected.

[94] (ʃ) is represented orthographically by "sh", which does not occur in Spanish; the borrowings in the corpus occur in words such as "champagne", "chef" and the like, so they are considered but not analysed in the following variable.

[95] The other 8 lexical items examined, plus the category *others* average N=40.6

[96] During the analysis, English was examined individually; the results show that the use of the target form was near-categorical for most speakers, which provides another rationale to limit the amount of tokens.

4.1.3 Variants and distribution

4.1.3.1 Variants

When the speech of the eighteen Chilean speakers of EFL of the sample is analysed, the data reveals that voiceless alveolar fricative (ʃ) presents four variants: one target and three non-target forms, as in the examples below, in this case represented orthographically by "sh" in the word *English*:

(12) a. "um, I like *English* ['ɪŋ.glɪ ᵗʃ], and I thought first to study medicine but, I thought, why not *English* ['ɪŋ.glɪʃ], *English* ['ɪŋ.glɪʧ] is, it's great, and I don't know, I like the possibilities that *English* ['ɪŋ.glɪʃ] offers." (Paola, female, first year speaker).

b. "my favourite subject was *English* ['ɪŋ.glɪt], I get good mark in *English* ['ɪŋ.glɪʃ], and history, I like it too." (Jaime, male, first year speaker).

The four variants were identified and defined as:

- *Voiceless alveolar fricative [ʃ]*, which corresponds to the target form in English; as mentioned earlier, this variant is stigmatised in Chilean Spanish. Hereafter, I will refer to this variant as "fricative".
- *Voiceless alveolar affricate [ʧ],* non-target form, corresponding to the Chilean prestige variant. This will be referred to as "affricate".
- *Voiceless alveolar plosive [t],* non-target form, to which I will refer to as "plosive"; this variant has not yet been shown in the literature.
- A form of *voiceless alveolar affricate with emphasis on the frication [ᵗʃ]*, non-target form; as the plosive, this has been shown briefly in the literature, as one of the variants identified by Cedergren (1973 – this is reviewed later).[97] From now on, this variant will be referred to as "fricativised". This form is particularly interesting, as it constitutes a "fudge" form in which the affricate and the fricative share at least one of its constituents (for more on fudge forms refer to e.g., Chambers and Trudgill, 1998). These fudge forms

[97] Cedergren (1973: 66) describes this form as a voiceless palatal affricate with a reduced stop onset. This feature accounts for the 8% of the total instances in her study.

usually arise from the contact between two languages or dialects (Chambers and Trudgill, 1980). In the case of this research, the use of the fudge form [tʃ] originates from the contact between Spanish (L1) and English (L2).

Despite the limited amount of studies among native speakers of English on voiceless alveolar fricative (ʃ), my research is focused on determining the source and patterns of variation in non-native speakers, where the feature under study does show systematic variation, conditioned by linguistic and social constraints. Specifically, I seek to test whether the patterns uncovered by Valencia (1983), Valdivieso (1983), Tassara (1992) and Cedergren (1973)[98] for native speech are also replicated in EFL contexts.

The following sections present this feature as studied previously and analyse its use on my sample, based on the social (individual speakers, sex, and years of instruction in English) and linguistic (preceding and following phonetic environment, lexical item, and position within the word) factors already introduced.

4.1.3.2 Distribution

4.1.3.2.1 Distributional data description for (ʃ)

The analysis of (ʃ) is presented in this section, according to the following social and linguistic factors: the effect of individual speakers, sex, years of instruction in English, socioeconomic group, lexical item, preceding and following phonetic environment, and position within the word.

4.1.3.2.1.1 Overall distribution

The overall distribution (henceforth OD) shows the "relative frequency of each variant of the variable without consideration of anything else" (Tagliamonte, 2006: 135). Thus, Table 4.1.1 shows the frequency of use for the four different alternative

[98] The sociolinguistic value of the feature under study in Chile opposes that of Panama City, where voiceless alveolar fricative is prestigious.

forms for the voiceless alveolar fricative (ʃ): fricative [ʃ], affricate [ʧ], plosive [t] and fricativised [ᵗʃ].

Table 4.1.1 Overall distribution for voiceless alveolar fricative (ʃ), all variants included (N=535)

	[ʃ]	[ʧ]	[ᵗʃ]	[t]	Total
%	83	11	5	1	
N	446	59	27	3	535

Table 4.1.1 shows that the fricative [ʃ], i.e., the target variant, accounts for 83% of the instances, followed by the affricate [ʧ] (11%), the fricativised [ᵗʃ] (5%) and finally the plosive [t] (1%).

Table 4.1.2 presents the results by individual speaker. The identification of the individual characteristics of speakers is essential to understand how they relate to (and differ from) the other members of the speech community (Table 4.1.2 arranges the speakers in descending order from the highest to the lowest percentage of use of the target form).

Table 4.1.2 Distribution of voiceless alveolar fricative (ʃ) all speakers included, all variants (N= 535)

	[ʃ]		[ʧ]		[ᵗʃ]		[t]	
All Speakers	%	N	%	N	%	N	%	N
Francisco	**100**	**22**	**0**	**0**	**0**	**0**	**0**	**0**
Bernardo	**100**	**24**	**0**	**0**	**0**	**0**	**0**	**0**
Miguel	**100**	**16**	**0**	**0**	**0**	**0**	**0**	**0**
Melinda	**100**	**30**	**0**	**0**	**0**	**0**	**0**	**0**
Caterina	**100**	**12**	**0**	**0**	**0**	**0**	**0**	**0**
Josefina	**100**	**49**	**0**	**0**	**0**	**0**	**0**	**0**
Carola	**99**	**90**	**1**	**1**	**0**	**0**	**0**	**0**
Carmen	**93**	**13**	**7**	**1**	**0**	**0**	**0**	**0**
Andres	**90**	**28**	**0**	**0**	**10**	**3**	**0**	**0**
Cesar	87	34	8	3	5	2	0	0
Cecilia	83	25	7	2	10	3	0	0
Juan	72	36	28	14	0	0	0	0
Consuelo	72	13	28	5	0	0	0	0
Marcela	60	6	40	4	0	0	0	0
David	58	11	5	1	37	7	0	0
Sebastian	57	12	43	9	0	0	0	0
Paola	48	12	24	6	28	7	0	0
Jaime	38	13	38	13	15	5	9	3
Total	**84**	**446**	**11**	**59**	**5**	**27**	**1**	**3**

The table shows that half of the speakers are (near-) categorical in their use of the target variant: four males (Francisco, Bernardo, Miguel, and Andres) and five females (Melinda Caterina, Josefina, Carola, and Carmen) present over 90% of use of the standard variant [ʃ];[99] these categorical speakers are shown in bold above. As mentioned in Section 3.4.3.2.4, the analysis of the individuals helps us identify how, and under which circumstances, the speakers vary. The information drawn from the analysis of individual speakers should determine how individuals contribute to the variation of the group; that is, which speakers behave similarly, but most importantly, who deviates from the group.

The first and most relevant finding is that nine speakers (the categorical and near-categorical) do not transfer the sociolinguistic patterns attested for Spanish; this means that they use the English target form /ʃ/ instead of avoiding it, despite this being stigmatised in their L1. Of the nine categorical speakers, six of them have 100% use of the target form (Francisco, Bernardo, Miguel, Melinda, Caterina and Josefina): Francisco is a first-year male of lower-middle class, Bernardo is a third-year male of middle-middle class, Miguel is a fifth-year male of upper-middle class, Melinda is a fifth-year female from the middle-middle class, Caterina and Josefina are third-year females from lower-middle and middle-middle classes, respectively; of all categorical speakers, only Josefina has attended English lessons outside this specific formal instructional setting.

All three near-categorical speakers are from fifth year; Carola and Andres are from the middle-middle class and Carmen from the lower-middle class.

Most female categorical speakers are concentrated in higher courses: Caterina and Josefina from third year and Carmen, Carola and Melinda all from fifth year, which suggests that the amount of time the informants have been exposed to English could be a highly (if not the most) relevant contributing factor in the use of the target form. In relation to class, lower-middle class presents 97% of use of the target, whereas middle-middle class and upper-middle class 98%. Since these differences are so

[99] Near categorical speakers are often considered those who present ≥95% of use of a particular variant (see, e.g., Tagliamonte, 2006). Here, I consider as near categorical speakers Carola and Carmen, as they show only 1 instance of a non-target form (99% and 93%, respectively), and Andres, who shows only 3 instances of a non-target form (90%).

minimal in relation to class, the variation due to exposure may be a more appropriate explanatory factor.

As mentioned above, the only one who has attended English lessons outside secondary school or university is Josefina, who shows a 100% (N=49) of use of the categorical [ʃ]; despite her having more years of instruction in English, her behaviour does not seem to differ from those other categorical speakers who have only attended the English teaching programme.

The use of the non-target variants by non-variable speakers is lower than the overall distribution: [ʧ] = 0.9%, and [ᵗʃ] = 1.1%. However, given the very small amount of tokens (N=18, 3% of the total), these results are far from being statistically significant.

Thus, despite the fact that the target form is stigmatised in Spanish, and, furthermore, that the voiceless alveolar fricative is not considered a phoneme in Standard Spanish, half the speakers use the target form categorically; therefore, the apparent most important constraining factor in the categorical behaviour of the above informants is years of experience in English, as the progression of categorical speakers is from one in first year, to three in third year and five in fifth year. Given the characteristics of the varying instances, particularly the small number of non-target forms, it is very difficult to test for other sources of variation or for the interaction of the different factor groups.

Following Tagliamonte's reasoning (2006: 86-87) in relation to the removal of categorical contexts ("the variable is realised either 0 per cent or 100 per cent of the time"), all (near) categorical speakers are removed from the data for two reasons: first, they obviously do not vary, so it is of no consequence to look for reasons of variation, and second, the inclusion of their data might alter the results of the other speakers whose use of the selected feature is, indeed, variable.[100] I remove also the variant [t], as it accounts for only 1% of the occurrences. Table 4.1.3 below shows the three variants left.

[100] See also Smith, Durham and Fortune (2007) and Smith, Durham and Richards (2013) for similar reasoning on the removal of (near) categorical speakers.

Table 4.1.3. Distribution of (ʃ) the variants [ʃ], [ʧ] and [ᵗʃ], non-categorical speakers (N=243)

	[ʃ]		[ʧ]		[ᵗʃ]	
	%	N	%	N	%	N
Total	67	162	24	57	10	24

Thus, the variation of 243 tokens is examined in relation to the selected constraining factors. This analysis is presented in the following section.

4.1.3.3 Factor by factor analysis

4.1.3.3.1 Social constraints

In order to test the effect of external factors on the variable use of voiceless alveolar fricative /ʃ/, and based on previous variationist research, I analyse four sets of factors: the effect of individual speakers, sex, years of instruction in English, and socioeconomic background. The effects of sex, years of instruction and socioeconomic background are particularly interesting to look at, given (1) the historical role of gender in variationist research, in which men and women have shown marked differences on linguistic variation, (2) the amount of exposure to English in formal instructional settings which might result in the increasing use of target forms (it is expected that the time informants have been exposed to English is concordant with the use of target forms), and (3) that the effect of social class has been shown to be an important contributing factor in variationist research.

4.1.3.3.1.1 Individual speakers

The distribution of voiceless alveolar fricative (ʃ) in non-categorical speakers is shown in Table 4.1.4 below. The preferred variant per speaker is shown in bold:

Table 4.1.4 Distribution of voiceless alveolar fricative (ʃ), non-categorical speakers only (N= 243)

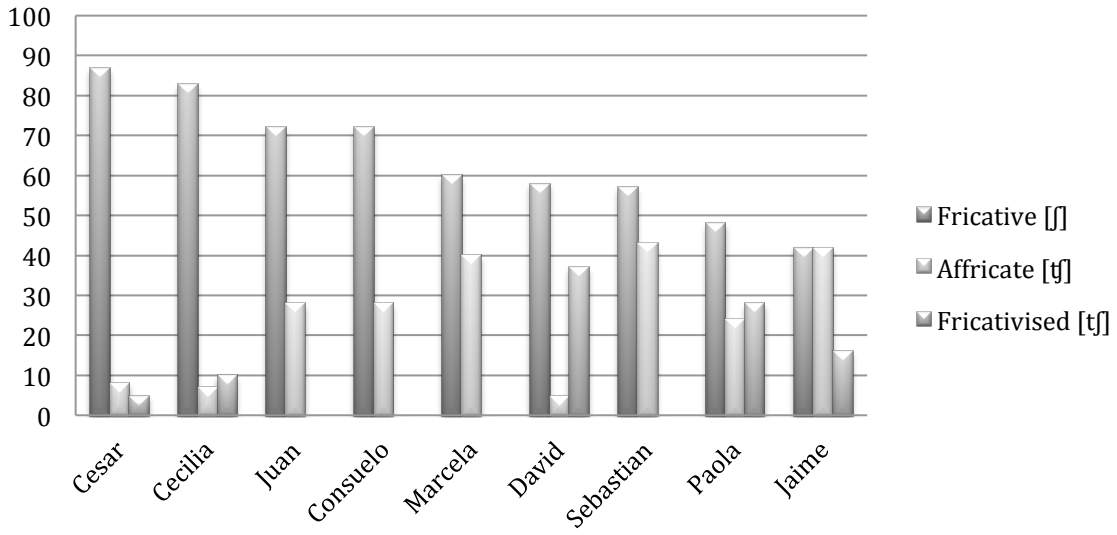
	Fricative [ʃ]		Affricate [ʧ]		Fricativised [ᵗʃ]	
Speaker	%	N	%	N	%	N
Cesar	**87**	34	8	3	5	2
Cecilia	**83**	25	7	2	10	3
Juan	**72**	36	28	14	0	0
Consuelo	**72**	13	28	5	0	0
Marcela	**60**	6	40	4	0	0
David	**58**	11	5	1	37	7
Sebastian	**57**	12	43	9	0	0
Paola	**48**	12	24	6	28	7
Jaime	**42**	13	**42**	13	16	5
Total	**67**	**162**	**24**	**57**	**10**	**24**

The predominant variant in almost every non-categorical speaker is the target form [ʃ], whereas non-target forms are noticeably less used. If we consider that the target form (L2) is the fricative, the developmental form the fricativised and the L1 form the affricate, a progression hierarchy should follow the pattern fricative > fricativised > affricate;[101] nonetheless, an analysis of the hierarchy of use suggests the following uses:

- Fricative > affricate > fricativised: Cesar
- Fricative > fricativised > affricate: Cecilia, David, Paola
- Fricative > affricate: Juan, Consuelo, Marcela, Sebastian
- Fricative = affricate > fricativised: Jaime

[101] This shall be considered further in Section 4.1.3.3.1.3, *Years of instruction in English*.

Figure 4.1.1 Distribution of voiceless alveolar fricative (ʃ), non-categorical speakers only (N=243)



The preliminary examination of the variation of (sh) suggests that:

- All nine non-categorical speakers use the target variant over the non-target variant, with the exception of Jaime, who uses the fricative and the affricate equally (42%, N=13).
- Four out of nine non-categorical speakers use the three variants, following the developmental hierarchy (fricative > fricativised > affricate): Cecilia and Paola from first year, and David, from third year (all three from different social groups). Of these, the use of the fricativised is particularly high in David (37%), however in general terms his behaviour is rather similar to Paola.
- Four out of non-categorical speakers use only fricatives and affricates, skipping the developmental variant: Juan (third-year male, upper-middle class), Consuelo (third-year female, middle-middle class), Marcela (first-year female, middle-middle class), and Sebastian (first-year male, lower-middle class).
- Two speakers stand out of the group: Cesar (fifth-year male, lower-middle class), as he exhibits a non-developmental hierarchy fricative > affricate > fricativised, which is dissimilar to the behaviour exhibited by other fifth-year speakers (almost all of them are categorical); and Jaime (first-year male,

lower-middle class), who shows the same percentage of use for fricatives and affricates; he also exhibits a relatively high use of the fricativised (>OD).

As patterns seem diffused across gender and social class, the number of years of instruction in English seems to be the clearest pattern so far to emerge: five speakers belong to first year, three to third year, and one to fifth year. A closer examination of all factors will reveal to what extent the instruction in English is playing a role in the variation of this group of non-native speakers.

4.1.3.3.1.2 Speakers' sex

The preliminary analysis revealed that the difference in variation does not seem particularly constrained by speakers' sex, as the use of both target and non-target forms is fairly similar for both sexes, as shown in Table 4.1.5.

Table 4.1.5 Distribution of voiceless alveolar fricative (ʃ) in all variable speakers, arranged by sex (N=243)

	Fricative [ʃ]		Affricate [ʧ]		Fricativised [ˢʃ]	
Speakers	%	N	%	N	%	N
Males	66	106	25	40	8	14
Females	68	56	21	17	12	10
Total	**67**	**162**	**24**	**57**	**10**	**24**

Intra-group analysis reveals a rather non-homogeneous behaviour, particularly amongst females. A marked difference in the use of the standard form for women over men was expected – in this case either the target fricative or the L1 standard affricate – but the difference was non-significant [102] (*p-value=0.5799272; Chi-square = 1.09*).

[102] Chi-squared tests are commonly used to test significance in linguistics. This test is used throughout this thesis, despite the fact other tests are preferred with contingency tables that contain small numbers (the criterion commonly used is less than 5 instances, which is not the case here). The test is carried out using the UCREL significance test system, an online Chi-square calculator provided by the University Centre for Computer Corpus Research on Language, from Lancaster University (http://corpora.lancs.ac.uk/sigtest/#extraHelp)

4.1.3.3.1.3 Years of instruction in English

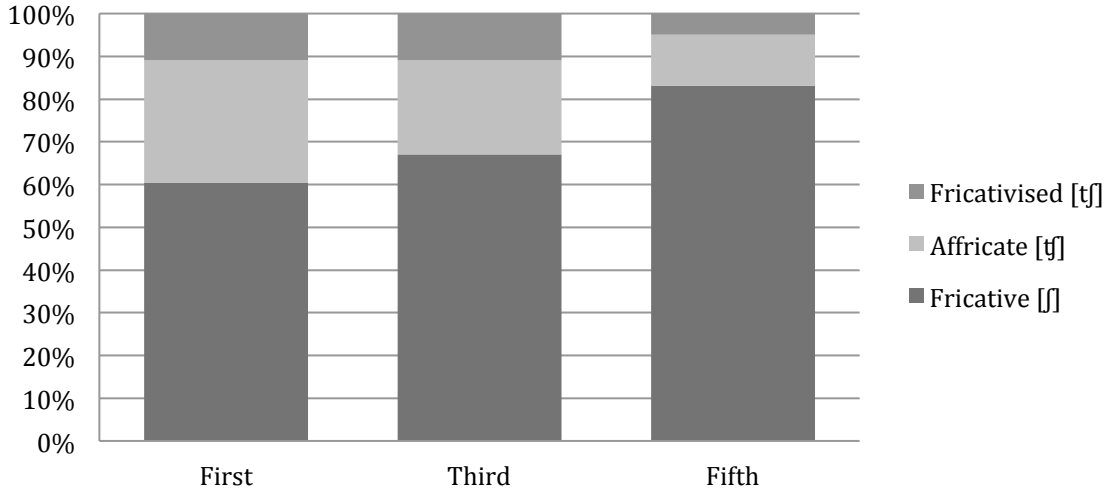
Table 4.1.6 shows that there is an increase in use of the target form from first to fifth year. This difference is statistically non-significant (*p-value=0.135824; Chi-square = 7.00*).

4.1.6 Distribution of voiceless alveolar fricative (ʃ) in all variable speakers, arranged by years of instruction (YIE) in English (N=243)

	Fricative [ʃ]		Affricate [ʧ]		Fricativised [ᵗʃ]	
YIE	%	N	%	N	%	N
Fifth	83	34	12	5	5	2
Third	67	60	22	20	11	10
First	61	68	29	32	11	12
Total	**67**	**162**	**24**	**57**	**10**	**24**

This progression could be explained by the fact that most fifth year speakers were removed as categorical and in this non-variant group only one of the fifth year individuals remains. The following graph displays more clearly how the three variants are distributed in relation to years of instruction in English.

Figure 4.1.2 Distribution of voiceless alveolar fricative (ʃ) in all variable speakers, arranged by years of instruction in English (N=243)



Despite the fact that the progression towards the use of the target fricativised from the non-target affricate seems to be clear,[103] intra-group analysis revealed the first and third years exhibited dissimilar behaviour, with a few using the developmental pattern suggested before, and others not; additionally, the differences between the three groups of speakers are minimal, which leads me to think that the variation of voiceless alveolar fricative cannot be attributable to the time of exposure to English. Therefore, the effect of other factors is explored.

4.1.3.3.1.4 Socioeconomic group (SEG)

In relation to the standard variant [ʃ], the behaviour of the three groups differs slightly from each other, increasing from lower-middle class to upper-middle class; the progression in use of target forms does not seem to be mirrored by the non-target variants (see Table 4.1.9 below); the use of affricates is similar in upper-middle and lower-middle class and fricativised are used by middle-middle and lower-middle class only.

A preliminary intra-group examination of socioeconomic class revealed similar patterns of variation in middle-middle and lower-middle-class speakers (refer to Table 4.1.4 for non-categorical individual speakers). The following table considers the group variation:

Table 4.1.7 Distribution of voiceless alveolar fricative (ʃ) per speakers' socioeconomic class, non-categorical, non-exceptional speakers only (N= 243)

	Fricative [ʃ]		Affricate [ʧ]		Fricativised [ˢʃ]	
Social class	%	N	%	N	%	N
Middle-middle	76	44	19	11	5	3
Upper-middle	68	47	22	15	10	7
Lower-middle	61	71	27	31	12	14
Total	**67**	**162**	**24**	**57**	**10**	**24**

The variation tested to be statistically non-significant (*p-value = 0.0.3721335: Chi-square = 4.26*); however, the use of the three variants is interesting for different reasons:

[103] When it is considered that a higher use of non-target forms might occur at the earliest stages of the acquisition process, with its consequent abandonment as the exposure to the language increases.

- First, the fricative, as already mentioned, is highly stigmatised in the L1 of the speakers; here the pattern suggests that the lower-middle class shows the lowest use of it (61%, N=71), whereas the highest use of this variant is in the middle-middle class (76%). In this case, it seems that lower-middle-class speakers are avoiding the use of their L1 stigmatised variant.
- Second, both non-target variants mirror the pattern of the target variant.
- Third, on the opposite side of the fricative, the prestigious form in L1 is the affricate. This shows its highest use by lower-middle-class (27%) and the lowest use in middle-middle class (16%). Traditionally, hypercorrection has been linked to lower-middle-class speakers in L1 (see e.g. Labov, 1966, 1990) and in L2 situations (Beebe, 1980); here this pattern seems to be replicated.
- Finally, the use of the fricativised also follows the same pattern as the affricate (middle-middle > upper-middle > lower-middle class). If we consider this a developmental feature, we may argue that it follows the same principle of hypercorrection; in other words, we might expect that the speakers who are aware of the value of the feature in the L1 and are still in the process of acquiring the target form, they might choose the prestigious form over the stigmatised, which would explain the pattern found here.

Despite the fact that the results that compare the three groups seem straightforward, with a clear pattern of variation, the differences due to social class do not seem to account for the variation of (ʃ) by itself.

4.1.3.3.1.5 Summary of findings - social factors

The main finding drawn from the initial analysis of the variable reveals that there are at least 4 variants for (ʃ), therefore variation as in L1 (Spanish) occurs in L2 contexts, the most frequent variant being the voiceless alveolar fricative [ʃ], which corresponds to the target form in English, as predicted.

In the factor-by-factor analysis, and once categorical speakers were removed, the results showed:

Individual speakers

All speakers show preference for the target form, with the exception of Jaime, who uses fricatives and affricates equally (42% each). This group was not tested for significance, as individual variation, despite its importance, cannot be statistically examined in relation to other individuals' based on any other characteristic than the ones studied as external factors (sex, age, years of instruction in English), at least with the regression model used in this research (see e.g. Johnson, 2009).

Sex

Added to the fact that differences between males and females were non-significant (Section 4.1.3.3.1.2), the patterns found were obscured due to non-homogeneous intra-group variation, meaning that the pattern generally exhibited by males versus females in other L2 studies (e.g., Major, 2004) is not replicated in relation to this particular feature.

Years of instruction in English

The results suggest that the use of target forms is progressive as students move through their acquisitional progress, most of them reaching the target form at the end of fifth year, which is explained by the fact that most categorical and near-categorical speakers are from fifth and third year (see Table 4.1.6). Also, there seems to be a developmental pattern in the use of non-target forms; however, the differences in variation tested to be non-significant and the intra-group analysis revealed non-homogenous behaviour.

Socioeconomic group

However statistically non-significant, the results for socioeconomic group are interesting as they show that the pattern of variation exhibited by all three groups are, firstly, internally homogenous and, second, there seems to be a clear distinction in the use of the variants per group, especially in relation to the higher use of the prestigious form in comparison to the target form in lower-middle-class speakers; this would be considered a case of hypercorrection.

However important these results are, it is necessary to consider the interaction of linguistic constraints in the variation of (ʃ), thus internal factors are analysed next.

4.1.3.3.2 Linguistic constraints

The internal linguistic factors considered in the analysis of voiceless alveolar fricative are both preceding and following phonetic environment, position in prosodic structure (initial and final), and lexical item. Based on previous research, phonetic environment may be an important conditioning effect (see Section 3.4.3.2.5, or, e.g., Cedergren, 1973).

4.1.3.3.2.1 Lexical item

Over 38 lexical items with more than ten instances of occurrence were coded, but after the exclusion of categorical speakers, as justified above, the lexical items were reduced to nine, plus the category *others* that compiles all instances <10; the remaining words are: *show (shown, shows), short (shorter, shorts), relationship, finish (finishing), English, should, she, share (sharing), and Spanish*. The distribution is presented in Table 4.1.8; this arranges the use of the target variant in descending order.

4.1.8 Distribution of voiceless alveolar fricative (ʃ) per lexical item, non-categorical, non-exceptional speakers only (N= 243)

Lexical Item	Fricative [ʃ]		Affricate [ʧ]		Fricativised [ˢʃ]	
	%	N	%	N	%	N
Show, shown, shows	87	13	13	2	0	0
Short, shorter, shorts	86	12	14	2	0	0
Relationship	80	12	20	3	0	0
Finish, finishing	71	10	14	2	14	2
English	70	30	9	4	21	9
Should	64	7	9	1	27	3
She	61	43	37	26	3	2
Share, sharing	60	6	30	3	10	1
Others	60	12	35	7	5	1
Spanish	57	17	23	7	20	6
Total	**67**	**162**	**24**	**57**	**10**	**24**

As justified in Section 3.4.3.2.7, I examine the effect of function versus content words; the analysis of word class – and in general of lexical item – is based on two facts: first,

it has been shown that high frequency lexical items behave differently than low-frequency ones (see, e.g., Bell *et al.,* 2003) and second, there are some similarities in the nature of the chosen words, that is, the variable position, the class, the phonetic context in which the variable occurs, etc. A higher use of weaker forms (fricatives, by process of lenition) is expected in function words than in content words, which will be strengthened by the effect of the plosive component of affricates.

Most lexical items are open class words,[104] with the exception of *should* (5% of the tokens, N=11) and *she* (29% of the tokens, N=71). The variation of these two words is similar in the target feature, but not in the non-target ones. *Should* presents a developmental hierarchy, with the highest use of the fricativised variant in the data (similar to that shown by the open class word *English*), whereas in *she* the use of the affricate is higher than in the other lexical items examined; just as *should*, *she* is rather similar to some open-class words, if compared, for instance, to the category *others*.

In sum, the class of the words seems to play no role in the variation, just as, tentatively, the position of the variable in the word (e.g., if we compare the groups *show* and *share*, the variable is in initial position, but the variation is dissimilar), and other structural characteristics – as in the lexical items *English* and *Spanish*, which not only present the variable in final position, but are also proper nouns, with the same preceding segment. So, how can such a difference be explained? A closer examination of this lexical item in interaction with other factors such as preceding and following phonetic segments may suggest an answer.

The examination of other factors is then required. The next factor to be considered is phonetic environment.

4.1.3.3.2.2 Phonetic environment

Historically, variationist studies of phonetic features have considered the observation of both previous and following segments as one of the most robust constraints that contribute to variation[105] (see for instance the contribution of phonetic environment in variables such as (t,d) deletion in Guy, 1994; Labov, 1997; Tagliamonte, 1998;

[104] The category *others* is made up exclusively of open class words.

[105] Labov (1989: 90) suggests that this is not a particularly strong constraint.

Tagliamonte and Temple, 2005; Bayley, 1994; Santa Ana, 1996; (s) deletion in Puerto Rican Spanish, Poplack, 1980). This section examines the effect of segment on the variation of voiceless alveolar fricative.

In the case of this variable, the essential study, Cedergren (1973), confirms the hypothesis that "the presence of a preceding vowel favours lenition", i.e., vowels, as preceding segment, favour the use of fricatives over affricates. In my research I coded all preceding and following sounds individually (cf. Appendix E for a detailed account on coding procedures and examples), including zero preceding or following segments as in the case of long pauses, initial or final occurrence of the feature; to facilitate the analysis of the effect of this factor group on the use of (ʃ), the individual segments were later collapsed into manner of articulation for consonants and vowels.

The results of the following analysis are dependent on the nine lexical items that result from the exclusion of all the categorical speakers,[106] plus the group *others*, therefore I will display the results for both the preceding and following segment, but the effect of these on the variation will be tested in interaction with the results drawn from the examination of lexical item.

4.1.3.3.2.2.1 Preceding phonetic environment (PPhE)

All preceding phonological instances occurring immediately before the variable were grouped in larger categories to facilitate the analysis;[107] this decision is supported by previous works such as Cedergren (1973),[108] Guy (1991), Poplack *et al.*, (2000a), Poplack and Tagliamonte (2001), among others. Guy (2007) suggests that the nature of the research objectives will determine how phonetic environment will be considered (and grouped), that is, it will depend on the level of analysis the researcher wants to conduct. Many studies consider the classification of all fricative consonants as fricatives, all plosive consonants as plosives, and so on, as shown in the examples below for plosives and fricatives.

[106] Most categorical items were produced by categorical speakers, thus the removal of these reduced the amount of lexical items to nine, as mentioned above.

[107] I provide an example of each preceding and following segment in Appendix E, for all variables examined.

[108] Cedergren (1973) goes even further and classifies the phonetic environment into three groups: vowels, consonants and pause.

Plosives:

- Plosive Alveolar /t/: "I try to *just short* objectives, reachable…"
- Voiced alveolar plosive /d/: "(...) from Europe *and she* knew Spanish."

Fricatives

- Voiceless alveolar fricative /s/: "that *is shorter*, but is one month I also…"
- Voiceless labiodental fricative /f/: "reinvented *herself, she* tried to make you…"

As the categories differ from each other, it is not possible to considered even larger categories, such as the comparison of vowels vs. fricatives seen in Cedergren, (1973) (for the comparison of vowels vs. consonants, or even according to the scale of sonority proposed by Major (1996: 97), see Section 2.1.3.2.2.1). The results are shown in Table 4.1.9.

Table 4.1.9 Distribution of voiceless alveolar fricative (ʃ) per preceding phonetic environment (PPhE): pause, vowels, and consonants; non-categorical, non-exceptional speakers only (N= 243)

	Fricative [ʃ]		Affricate [ʧ]		Fricativised [ᵗʃ]	
PPhE	%	N	%	N	%	N
Fricatives	87	13	7	1	7	1
Laterals	86	6	14	1	0	0
Pause	83	5	17	1	0	0
Nasals	70	19	30	8	0	0
Vowels	67	99	20	29	14	20
Tap	67	6	33	3	0	0
Plosives	45	14	45	14	10	3
Total	**67**	**162**	**24**	**57**	**10**	**24**

The differences in the variation for this group tested to be significant (*p-value = 0.000006393875, Chi-square = 46.21).* However, as suggested, e.g., in Cedergren (1973), vowels do not seem to favour the use of fricatives more than, in general, obstruents. Furthermore, recall that *English* and *Spanish*, both with the same preceding segment, presented different behaviour, with one of them showing a noticeable higher use of fricatives than the other.

On the other side, it is interesting to notice that plosives do seem to show a higher use in affricates and fricativised (45% and 10%, respectively), considering that both have a plosive component in their articulation, when compared with fricatives.

Thus, preceding segment, even though it is significant, does not seem to constrain the variation of voiceless alveolar fricative on its own, therefore the study of following phonetic environment is required.

4.1.3.3.2.2.2 Following phonetic environment (FPhE)

As with preceding phonetic environment, all following segments examined (after the removal of categorical speakers) are grouped into larger categories, as justified before.[109] The following examples illustrate the categories nasals and vowels:

Nasals

- Bilabial nasal /m/: (…) "as a *punishment*, I mean."
- Alveolar nasal /n/: "a very good *English knowledge* but…"

Vowels

- Close front (short) vowel /ɪ/: "(…) said that *English is* the most important…"
- Back close-mid vowel /o/: "is a kind of *show* in the park there…"

The same hypotheses should apply here: vowels favour the use of fricatives and plosives of affricates and fricativised. The following table introduces the results; it is arranged showing the target form in descending order.

[109] As with preceding segment, examples of all found environments are provided in Appendix E.

Table 4.1.10 Distribution of voiceless alveolar fricative (ʃ) per following phonetic environment (FPhE); non-categorical, non-exceptional speakers only (N= 243)

	Fricative [ʃ]		Affricate [ʧ]		Fricativised [ˢʃ]	
FPhE	%	N	%	N	%	N
Nasals	100	5	0	0	0	0
Approximants	100	3	0	0	0	0
Affricates	100	1	0	0	0	0
Tap	100	1	0	0	0	0
Pause	70	9	8	1	23	3
Vowels	67	122	28	51	6	10
Plosives	61	14	17	4	22	5
Fricatives	50	7	7	1	43	6
Total	**67**	**162**	**24**	**57**	**10**	**24**

The results are statistically significant (*p-value = 0.001395195; Chi-square = 35.15*). Based on the factor-by-factor analysis, in which preceding and following segments are analysed independently, the hypothesis that vowels would favour the use of the fricative is confirmed, but consonants do not seem to restrain its use, as argued by Cedergren (1973), at least with following phonetic environment. Furthermore, fricatives show the opposite pattern as preceding segment (as following, they show the same percentage of use for target and non-target variants). Only vowels and plosives seem consistent, vowels with similar percentages and hierarchies of use,[110] and plosives with a relatively low percentage of use of targets in comparison with non-target variants.

Before proceeding to the analysis of the interaction of factors, I examine the position of the variable within the word.

4.1.3.3.2.3 Position within the word

Prosodic position has been determined as a powerful contributing factor in the study of, for instance, consonant strengthening, in which initial position tends to favour the strengthening and medial position the fricativisation, especially intervocalically (see e.g., Fougeron and Keating, 1997; Fougeron, 2001; Pérez, 2001, 2007). Table 4.1.11 shows the distribution of lexical item in relation to initial, middle and final positions.

[110] A cross tabulation of preceding and following segment confirmed that intervocalic position favoured fricatives (67%, N=60), which is consistent with the patterns found by Cedergren (1973).

Table 4.1.11 Distribution of voiceless alveolar fricative (ʃ) per position within the word; non-categorical, non-exceptional speakers only (N= 243)

	Fricative [ʃ]		Affricate [ʧ]		Fricativised [ˢʃ]	
Position	%	N	%	N	%	N
Middle	**77**	**20**	23	6	0	0
Initial	67	84	**28**	**35**	5	6
Final	63	58	17	16	20	18
Total	**67**	**162**	**24**	**57**	**10**	**24**

Firstly, the effect of prosodic position tested to be significant (*p-value 0.001402926; Chi-square = 17.71*). Second, the results seem consistent with the literature, as in initial position the percentage of affricates is higher than in middle position (28%, shown in bold); accordingly, the highest percentage of fricatives occurs in middle position (77%, in bold). More interestingly, final position shows a relatively high percentage of use of non-target forms. The articulation of affricates (and of consonants in general) in the L1 of the speakers is not as common as in English (see e.g. Alvar, 1980); even less common is the kind of features that is been articulated, in this case, both affricates and fricativised have a plosive and a fricative component, which rarely occurs in the L1 of the learners. This means that the learners have not only acquired the syllabic structure of the L2 (in which the articulation of consonants is more frequent in coda position), but they are using the prestigious and the fudge (developmental) forms over the target form (17% and 20%, respectively, against 63% of fricatives).

A closer examination to the variation of each lexical item in relation to the prosodic position of the variable would give us more information on how these two factors interact. This interaction is shown in the cross-tabulation, presented in Table 4.1.12 below.

Table 4.1.12 Cross tabulation of individual lexical item and position (initial, middle and final) (N=243)

	Lexical Item	Variant	N	%
Initial position	Show, shown, shows	Fricative	13	87
		Affricate	2	13
		Fricativised	0	0
		Σ	15	
	Short, shorter, shorts	Fricative	12	86
		Affricate	2	14
		Fricativised	0	0
		Σ	14	
	Others	Fricative	3	75
		Affricate	1	25
		Fricativised	0	0
		Σ	4	
	Should	Fricative	7	64
		Affricate	1	9
		Fricativised	3	27
		Σ	11	
	She	Fricative	43	61
		Affricate	26	37
		Fricativised	2	3
		Σ	71	
	Share, sharing	Fricative	6	60
		Affricate	3	30
		Fricativised	1	10
		Σ	10	
Middle position	Finished, finishing	Fricative	4	100
		Affricate	0	0
		Fricativised	0	0
		Σ	4	
	Relationship	Fricative	12	80
		Affricate	3	20
		Fricativised	0	0
		Σ	15	
	Others	Fricative	4	57
		Affricate	3	43
		Fricativised	0	0
		Σ	7	
Final position	English	Fricative	30	70
		Affricate	4	9
		Fricativised	9	21
		Σ	43	
	Finish	Fricative	6	60
		Affricate	2	20
		Fricativised	2	20
		Σ	10	
	Spanish	Fricative	17	57
		Affricate	7	23
		Fricativised	6	20
		Σ	30	
	Others	Fricative	5	56
		Affricate	3	33
		Fricativised	1	11
		Σ	9	

The examination of the lexical items per position confirms that:

- All lexical items favour the use of the target fricative. The three variants are used in initial and final position, whereas only fricatives and affricates are used in middle position.
- The frequency of use of the target form decreases in the pattern fricative > affricate > fricativised for most lexical items, with the exception of should, which presents the developmental pattern fricative > fricativised > affricate.[111]
- Conversely, the use of non-target forms increases in middle and final positions; the use of affricates is higher for most lexical items with the exception of English, which shows the highest use of the fricativised variant (in final position).

In sum, the results indicate that the use of the target forms is higher in initial and middle positions and non-target forms are preferred in final position. This is partially consistent with the literature mentioned earlier, as studies have shown that stronger articulations occur in initial position (affricates, by effect of their plosive component), whereas middle position tends to favour fricativisation; the high percentage of non-target forms in final position suggests that the speakers have acquired the syllabic structure of English, but they might be still in the process of separating the sociolinguistic characteristics of this feature from their L1 and the L2.

4.1.3.3.2.4 Summary of findings: linguistic constraints

The analysis of the linguistic constraints reveals the following findings:

Lexical items

After the removal of categorical and exceptional speakers, the lexical items analysed are: *show (shown, shows), short (shorter, shorts), relationship, finish (finishing), English, should, she, share (sharing), and Spanish* plus a group that considers all words

[111] The pattern of variation of *should* is similar to the lexical item *English*. The relatively high variation of these two words is particularly interesting given the saliency of these words for learners of English, as high frequency words (Source: Corpus of Contemporary American English, http://www.wordfrequency.info).

with less than 10 occurrences. When examined as individual lexical items, the only pattern that arises is the high use of the fricatives over the other variants.

Phonetic environment

The phonetic environment in which the variable occurs is bound to the lexical item, thus the examination of segment is carried out to confirm this phenomenon. The results are statistically significant for preceding and following segment; however, their contribution to the variation of voiceless alveolar fricative is not marked as suggested in the literature, for each of the groups examined; furthermore, the effect of segment does not seem to be related to the lexical items that share similar structural characteristics (such as the words *English* and *Spanish)*.

Position within the word

In general terms, a preliminary examination of position of the variable within the word suggests that this is one of the most robust contributing factors, as the use of fricatives and affricates shows a more marked difference than with other factors, in middle and initial position, respectively; this tested to be significant. The interaction of position and lexical item seems to reveal that the effect of position, just as with preceding and following segments, is bound to the lexical item in which the variable occurs.

Having examined the variation of individual factors, I now turn to the multivariate analysis, which will show the contribution of all factor groups considered at the same time, to the variation of voiceless alveolar fricative.

4.1.3.3.3 Multivariate analysis

In order to conduct multivariate analysis, I examined the application of the rule on the target variant [ʃ] only, to avoid the occurrence of empty cells; thus, the combination of factors is observed at least once. In the factor group *Following phonetic environment* I removed the categorical segments, as these do not contribute to the examination of variation; this results in a reduction of the total number of tokens to N= 231.

The results are shown in Table 4.1.13 below; these are arranged from the highest to the lowest contributing factor and in each of them from the highest to the lowest factor weight. Six out of the eight factor groups are examined: from the social factors I included *SEG* and *years of instruction in English*, and from the linguistic ones I included *lexical item*, *preceding* and *following phonetic environment*, and *position*. The decision to examine only these factors is made to avoid the "kitchen sink effect" (Tagliamonte, 2006).

Table 4.1.13 Multivariate analysis for voiceless alveolar fricative (ʃ)

	Factor Weight (FW)	%	Total N
Input .654		65	231
PPhE			
Fricatives	.77	87	15
Laterals	.75	86	7
Pause	.72	83	6
Nasals	.55	70	27
Tap	.50	67	9
Vowels	.49	65	137
Plosives	.28	43	30
Range	**49**		
YIE			
Fifth Year	.71	83	41
Third year	.50	66	84
First year	.42	59	106
Range	**29**		
SEG			
Middle-middle	.61	75	56
Upper-middle	.51	67	63
Lower-middle	.44	60	112
Range	**17**		
FPhE			
Pause	[.54]	69	13
Vowels	[.51]	67	183
Pause	[.45]	61	23
Fricatives	[.35]	50	12
Lexical Item			
Show, shown, shows	[.77]	87	15
Short, shorter, shorts	[.75]	86	14
Relationship	[.67]	80	15
English	[.52]	68	41
Should	[.47]	64	11
She	[.43]	60	70
Finish, finishing	[.43]	60	10
Others	[.41]	59	29
Spanish	[.37]	54	26
Position			
Initial	[.51]	67	123
Final	[.44]	60	82
Middle	[.25]	77	26

As observed, three factor groups contribute to the variation of (ʃ): preceding phonetic environment (range 49), years of instruction in English (range 29), and socioeconomic group (range 17). The other three factor groups (following segment, lexical item and position) were not significant to the variation of voiceless alveolar fricative.

4.1.3.3.4 Summary and conclusions

The initial hypothesis postulated was that the use of voiceless alveolar fricative (ʃ) found in the literature for Chilean Spanish native speech would be replicated in English, which has been refuted: speakers use the English target form [ʃ] considerably more than the Chilean Spanish prestigious [ʧ], despite the fact that the target form is stigmatised in Chilean Spanish. This finding suggests that speakers overcome the social meaning of the feature and acknowledge the phonetic differences between the L1 (Spanish) and the L2 (English). However, the use of the non-target forms seems to be sociolinguistically motivated: both social and linguistic factors have been found to contribute to the variation of voiceless alveolar fricative.

Specifically, in relation to the hypotheses proposed, the multivariate analysis suggests:

- The most significant factor is preceding segment (range 49). As opposed to the results shown by Cedergren (1973), vowels do not disfavour the use of fricatives by effect of lenition; of this factor group, fricatives promote the use of the fricative variant the most (FW=.77) and plosives the least (FW=.28). The multivariate analysis demonstrates that this is the only internal factor that contributes to the variation of voiceless alveolar fricative (ʃ).
- The second factor found significant is years of instruction in English (range 29), showing a clear progression in the acquisition of the target form from first (FW=.42, 59% of target use) through third (FW=.50 66% of target use) to fifth year (FW=.71, 83% of target use).
- Finally, the third contributing factor is socioeconomic group (range 17). The results confirm the findings of Labov (1966, 1990) and Beebe (1980) which suggest that hypercorrection occurs in lower-middle class speakers; here, this group shows a relatively high use of the non-target variants (more

prestigious ones) if compared to the target form (FW=.44, 60%). The highest use of the target form occurs in middle-middle-class speakers FW=.61, 75%).

Of the tested hypotheses, the effect of years of instruction in English is one of the most relevant factors for the study of second language acquisition variation, as it shows that the acquisition of target forms is actually linked to exposure to the second language, even in foreign language settings. I turn to this point later in Chapter 5, in the discussion section for voiceless alveolar fricative. Now, I examine the results for voiceless alveolar affricate.

4.2 Voiceless alveolar affricate (tʃ)

4.2.1 The variable

Voiceless alveolar affricate (tʃ) occurs in English in a range of different positions and it is usually represented by the digraph "ch" in contexts as *chip* /tʃɪp/, *teacher* /ˋti:tʃər/, *bleach* /bli:tʃ/. The examples below show four alternative uses for voiceless alveolar affricate in the word *teacher*:

(13) a. "I had a teacher [ˋti:ʃər] from Canada"
(Cecilia, first year, female speaker)

b. "Most of my teacher [ˋti:tʃər] are very good teachers"
(Paola, first year, female speaker)

c. "I'd like to be a teacher [ˋti:tʃ̍ər]…"
(Josefina, thrid year, female speaker)

d. "…he was a good teacher [ˋti:ᵗʃər]".
(Carola, fifth year, female speaker)

In contrast to English, in Spanish voiceless alveolar affricate does not occur in final position; I will revisit the implications of position later, but as shown for the first variable examined, in Spanish, consonants in coda position tend to be aspirated, weakened or deleted (see e.g., Pérez, 2007 for voiced dental fricative). What is most interesting about this feature is that in Chilean Spanish – as with the first variable examined – the sociolinguistic value of this variable in Spanish suggests that this feature might be (1) used in a target-like fashion, (2) hypercorrected by means of the

use of a recent prestigious variant of Chilean Spanish (see Valencia, 1993-1994) or (3) replaced by a stigmatised feature (Valdivieso, 1983; Tassara, 1992; Valencia, 1993-1994), also present in the previous variable.

In an attempt to uncover how students use (ʧ) according to a number of linguistic and social correlates, and how they behave when faced with a feature that has different sociolinguistic values in Chilean Spanish and in English, in this chapter I will discuss the use of the variable selected, including the sociolinguistic value it has in Chilean Spanish, and I will then show the factors found to be contributing to the use of (ʧ), which will be followed by the analysis and discussion of the actual presence of the variable in my sample.

4.2.2 Previous findings and specific hypotheses

Just like with the previous variant, voiceless alveolar affricate (ʧ) has not been studied in native speakers of English, but it has been the object of variationist research in English as L2 and in Spanish as L1. Furthermore, given that this variable is considered one of the variants for the first variable under examination, the same SLA research – and to some extent, the same literature on Chilean Spanish – covers voiceless alveolar fricative (ʃ) and voiceless alveolar affricate (ʧ). The following section reviews some studies that observe voiceless alveolar affricate directly as a variable.

4.2.2.1 Variation in Spanish

The variation of voiceless alveolar affricate has been noted in Chilean Spanish by Salamanca and Marlet (2001), who refer to the alternation of (ʧ) in Chilean Spanish in two realisations: [ʧ] and [ʃ]. However, Valencia (1993-1994) shows that (ʧ) presents three variants in the speech of Chilean Spanish: [ʧ] as the affricate variant, [ʃ] as the fricative variant and [tʃ] as an innovative variant, described as a very tense plosive followed by a very brief friction (fricative – Valencia, 1993-1994: 168).[112] From Valencia´s study, and the contributions of the theory in relation to gender and middle-class sociolinguistic behaviour, we can assume that in Chilean Spanish there are at

[112] My own translation of "una plosiva muy tensa, seguida de una fricción muy breve" (Valencia, 1993-1994: 168).

least three sociolinguistic values for the variable (ʧ): a *standard* variant that keeps a relatively stable distribution in all three social classes, independent of age and gender, and is represented by the affricate variant [ʧ]; a *stigmatised* variant, present mostly in the lower social class and older speakers, represented by the fricative variant [ʃ];[113] and a *prestigious* variant, used mostly by female, younger speakers from the middle social class, represented by the innovative variant [tʃ]. These three variants are present in my sample, as we will see in the following sections. Each of the factors mentioned above are detailed next.

In her study, Valencia (1993-1994: 160) examines the use of the phoneme (ʧ) in the informal speech of 24 adolescents from Santiago; the factor groups considered are age (14-17 years old), sex and "socioeconomic and cultural level" (index made up of level of education of the individuals, profession and/or education and income of parents). In sum, Valencia's findings suggest that:

- The use of the innovative variant depends on factors such as sex – female speakers showed the highest percentage of use (28.8%), and the highest difference in relation to male speakers' use of the same variant (21.7% of difference).
- The social class of speakers is a relevant contributing factor with a clear distinction between standard and stigmatised variants in the three socioeconomic groups studied (upper, middle and lower).

These findings are supported by the works by Cedergren (for Panama Spanish, 1973), Bernales (1978), Widgorsky (1978), Valdivieso (1983, 1993), Tassara (1992), Valencia (1993), and Cepeda (2001).

4.2.2.2 Variation in ESL/EFL contexts

As with the previous variable, the variation of voiceless alveolar affricate has been linked to different communities of speakers of English as a second or foreign language, predominantly in connection with a Latino background (Spanish L1

[113] In Cedergren (1973, 1987) the use of the fricative was mostly used by middle-aged, middle-class, female speakers, suggesting, according to the literature, that the use of the fricative variant was considered prestigious, just the opposite as in Chilean Spanish.

speakers). As this variable has not been examined as such, but in relation to the alternation (sh/ch) or as an approximation to (s) in several dialects, the same literature mentioned in Section 4.1.2 covers the variation of (tʃ) in non-native English, for both immigrant and instructional settings. Now I turn to the specific hypotheses formulated for this variable.

4.2.2.3 Hypotheses

The systematic variation of voiceless alveolar affricate (ʧ) evidenced in native speakers of Spanish is partly transferred to non-native English by means of the use of Chilean Spanish prestigious forms; that is, it is expected that if variation occurs, it can be justified on the grounds of hypercorrection[114] by, e.g., the use of [tʃ], which has not been documented in English.[115]

4.2.2.4.1 Specific hypotheses

As justified in Section 3.1.2, and given that each variable is different to each other, it is necessary to formulate individual sets of hypotheses; these are justified by the reviewed literature (Sections 4.2.2.1 and 4.2.2.2). For voiceless alveolar affricate the secondary hypotheses are:

- High rates of the target variant are predicted, as [ʧ] also corresponds to the standard form in Chilean Spanish.
- The sex of speakers is determinant in the selection of the alternative forms, with female speakers favouring the prestigious and innovative forms, and males using the standard over the non-standard ones, as shown in Valencia (1993-1994).
- A progressive use of the canonical form – affricate – is expected according to the amount of exposure to English in formal instruction, with the consequent abandonment of the non-canonical forms.
- Socioeconomic class is a relevant constraint, especially in relation to the use of prestige versus standard features.

[114] Cf. Section 4.1.2.4, for voiceless alveolar fricative.

[115] To this author's knowledge.

- An important effect of the phonetic environment on the use of (ʧ). Given the results from voiceless alveolar fricative, a distinction on vowels and consonants is expected, with vowels favouring the less strongly articulated forms (e.g., fricativised). For prosodic position within the word, a higher use of affricates in initial position is expected, given the obstruent present in affricates; this tends to strengthen in initial positions and weaken intervocalically.

4.2.2.4 Exclusions

All the tokens containing the variable (ʧ) were extracted from the corpus. I searched through for all words containing "ch", as this is the corresponding orthographic realisation in both English and Spanish. Words in Spanish, as in examples 14a and 14b, and borrowings (as in *champagne, chef*) are excluded as the study focuses on the use of English forms only, and the pronunciation of, for instance, French borrowing differs from the pronunciation for English and Spanish. Words such as "culture" in which the variable also occur were not included, as I focused on the most common orthographic representation for (ʧ) – *ch* - whose pronunciation could be easily identified by Spanish speakers; despite being a considerable methodological drawback, I decided not to code for items like "culture" as they might result in higher rates of variation that might skew the results; however, this fact should be taken into consideration in further extensions of this research. Although the word *Chile* appears 56 times (7.7%), it will not be considered further for the same reasons here presented, i.e., its pronunciation might be affected by the L1 pronunciation. However, *Chilean* will be considered as it takes an English suffix (the same methodological consideration taken with the words *English* and *Spanish* in the previous variable),[116] and is morphologically different from the Spanish word.

(14) a. "I wanted to stay here in *Chillán*, so my, this was my, my first option."
Josefina (female, third year speaker)
b. "We eat a lot of empanadas, *chicha*, er, barbeques…"
Marcela (Female, first year speaker)

[116] The inclusion of words such as *Chilean, English, Spanish* and other similar lexical items responds primarily to the fact that they are English words; in the case of Chilean, this takes the nationality suffix –ean, that does not reflex the Spanish nationality suffix for the equivalent word (-eno) (it is not even considered a cognate or loanword). Thus, the exclusion of the word *Chile* is in keeping with the fact that the word is a proper name, with no translation or equivalent in English.

Having defined the variable contexts, I now turn to the results of the analysis.

4.2.3 Variants and distribution

4.2.3.1 Variants

After a preliminary examination of the variation of voiceless alveolar affricate, four variants were identified by means of auditory analysis. These are presented here:

- Voiceless alveolar fricative [ʃ]: corresponds to the Chilean stigmatised variant, henceforth *fricative.*
- Voiceless alveolar affricate [ʧ]: the standard form in English and in Spanish; from now on I will refer to it as *affricate*.
- A form of voiceless alveolar affricate with emphasis on the stop [tʃ]: this corresponds to the prestigious form of recent appearance in Chilean phonetic inventory, described by Valencia (1993-1994); it will be referred to as *stop*.[117]
- A form of voiceless alveolar affricate with emphasis on the frication [^{t}ʃ]: Cedergren (1973) refers to this variant as an intermediate form ("fudge" form, as in Section 4.1.3.1, for voiceless alveolar fricative) that occurs between the affricate [ʧ] and the fricative [ʃ] in the process of (ch) lenition in the speech of Panama City; from now on, this variant will be referred to as *fricativised*.

4.2.3.2 Distribution

Having introduced the four variants for (tʃ), I now introduce the results obtained from the analysis of the selected variable. First, I circumscribe the context for the variable – specifically, I select those contexts which should be excluded from the analysis. Then I move to the actual results (Section 4.2.3), beginning with the distributional data and then follow with the analysis of the individual factors. I end this section with a multivariate analysis and the corresponding conclusions.

[117] This form is not a stop, as such, but for practical purposes (easier identification) it will be referred to as stop, as it presents an increased obstruent onset, which means that the emphasis on the articulation is in the stop component of the fricative.

4.2.3.2.1 Distributional data description for /tʃ/

The following section introduces the results obtained from the preliminary analysis of the data and it is divided into two parts: the overall distribution of voiceless alveolar affricate and then the examination of the behaviour of categorical and exceptional speakers.

4.2.3.2.1.1 Overall distribution

The overall distribution shows the frequency of use of all four variants for (ʧ).

Table 4.2.1 Overall distribution of voiceless alveolar affricate (ʧ)

	[tʃ] Stop	[ʧ] Affricate	[tʃ] Fricativised	[ʃ] Fricative	Total
N	294	285	110	37	726
%	41	39	15	5	

This table reveals that the highest percentages of use occur in the stop [tʃ] (41%), followed very closely by the affricate [ʧ] (39%), which is consistent with the high rates of use of these two variants in Chilean Spanish found in Valencia (1993-1994). However, we must remember that the interviews were conducted in English, and so all the instances occurred in English words, thus we cannot apply Valencia's results to this study directly, but only take them as a reference of what occurs in Chilean Spanish. Of these, the target form for English is the variant affricate [ʧ]; of the non-target forms, the fricative [ʃ] is considered stigmatised in Chilean Spanish, but not in English; the stop, which is a fudge form consisting of an affricate with higher emphasis on the plosive (hence the notation assigned to it here, [tʃ]); finally, a second fudge form is present in the data, with a comparatively small amount of instances; this has been considered a transition or developmental sound and it is characterised as an affricate with emphasis on the fricative component [tʃ]. It also shows that the fricative [ʃ] is a minority variant accounting for only 5% of the data (N=37); as it occurs in very restricted contexts, it is henceforth excluded from the analysis.

What is relevant is the high use of [tʃ] over the standard [ʧ], against the initial prediction. This result is very interesting to observe, as the analysis will contrast the use of these two variants that constitute the standard and the prestigious forms for

English and Chilean Spanish, respectively, plus the other fudge form mentioned ([ᵗʃ]). A new total of 689 tokens are examined.

Table 4.2.2 Distribution of voiceless alveolar affricate (tʃ), all speakers, three majority variants (N=689)

	[ʧ] Affricate		[ᵗʃ] Stop		[ᵗʃ] Fricativised	
Speaker	%	N	%	N	%	N
Andres	71	15	0	0	29	6
Melinda	70	26	8	3	22	8
Cecilia	67	16	17	4	17	4
Bernardo	66	29	21	9	14	6
Francisco	61	14	30	7	9	2
Cesar	60	38	32	20	8	5
Miguel	56	28	14	7	30	15
Carola	42	53	17	22	41	52
Sebastian	41	13	59	19	0	0
Paola	35	8	52	12	13	3
Carmen	35	14	65	26	0	0
Consuelo	27	3	73	8	0	0
David	25	11	66	29	9	4
Caterina	22	5	78	18	0	0
Josefina	19	6	81	25	0	0
Jaime	12	3	76	19	12	3
Marcela	*7*	*1*	*87*	*13*	*7*	*1*
Juan	*4*	*2*	*95*	*53*	*2*	*1*
Total	**40**	**285**	**48**	**294**	**12**	**110**

The distribution of affricates and stops is similar (40% and 48%, respectively). There is only one near categorical speaker (shown in italics) with 95% of use of a single variant (Juan, 95% of the stop, N=53). Marcela (also shown in italics) can be considered categorical not because of the percentage of use of the stop, but because she uses the affricate only once.[118] The first overall finding suggests that, as opposed to the previous variable where half of the speakers were (near) categorical, the variation between speakers is robust; so the question is what factors can explain these differences in the variation? Will the effect of years of instruction and social class be as relevant as for voiceless alveolar fricative, or will class have an effect on its variation? I start by examining the two categorical speakers.

[118] Note that the numbers are, in general, very small, but this is because the interviews averaged 47 minutes, therefore a large number of tokens of this particular variable cannot be expected.

4.2.3.2.1.2 Categorical and exceptional speakers

Juan is a male speaker from third year who belongs to socioeconomic group C2. He has not attended English lessons outside the instructional setting. His use of the three variants is consistent with the hierarchy shown in the OD: stop > affricate > fricativised. The only lexical item in which he presents variation is the word *teacher*, and in all the others he uses the stop.[119] The other exceptional speaker, Marcela, a female first-year informant from group C3, has not attended English lessons outside the instructional setting. She favours the use of the stops over the affricate (92%, N=11), despite the small number of instances (N=12, total). Her varying instances occur in different lexical items, the most frequent being *child, childhood* (N=4, 80% of the stop), and *much* (N=3, 100% of the stop); the five remaining instances occur in five different words. As no similarities can be detected demographically, or in relation to lexical item, variation is attributed exclusively to individual differences. As the inclusion of these two speakers might skew the results, they are hereafter removed from the analysis. Thus, a total of 618 tokens are examined.

4.2.3.3 Factor-by-factor analysis

4.2.3.3.1 Social constraints

The effect of social factors is tested to determine the source of variation of voiceless alveolar affricate and the possible transfer of L1 features into English; thus, four sets of factors are examined: individual speakers, sex, years of instruction in English, and socioeconomic group (SEG); as with the previous variable, age is not considered given that the speakers belong to the same age group. The factors that are expected to exhibit greater effect on variation are sex and socioeconomic group, as (1) prestigious over standard features have been tested to be preferred by women, and (2) SEG has particularly shown great influence on the variation of this feature in Chilean Spanish and in variationist research in general.

[119] A closer examination of lexical items for all speakers is given in Section 4.2.3.3.2.1.

4.2.3.3.1.1 Individual speakers

The effect of individual speaker has been discussed before: it is expected that, beyond any possible pattern that might emerge from the different factor groups, the individual speakers might behave differently to each other; Table 4.2.3 below shows non-categorical speakers. The highest percentages per speaker are shown in bold; these are displayed from the highest to the lowest percentage of use of the preferred variant (affricate):

Table 4.2.3 Distribution of voiceless alveolar affricate (tʃ), non-categorical speakers (N=618)

	[ʧ] Affricate		[tˢ] Stop		[ˢʃ] Fricativised	
Speaker	%	N	%	N	%	N
Andres	**71**	15	0	0	29	6
Melinda	**70**	26	8	3	22	8
Cecilia	**67**	16	17	4	17	4
Bernardo	**66**	29	21	9	14	6
Francisco	**61**	14	30	7	9	2
Cesar	**60**	38	32	20	8	5
Miguel	**56**	28	14	7	30	15
Carola	**42**	53	17	22	41	52
Sebastian	41	13	**59**	19	0	0
Paola	35	8	**52**	12	13	3
Carmen	35	14	**65**	26	0	0
Consuelo	27	3	**73**	8	0	0
David	25	11	66	29	9	4
Caterina	22	5	78	18	0	0
Josefina	19	6	81	25	0	0
Jaime	12	3	76	19	12	3
Total	**46**	**282**	**37**	**228**	**18**	**108**

After the removal of the categorical speakers, the initial hierarchy shown in Table 4.2.1 has changed: affricates are more frequent than stops. The difference in the rates of use of affricates and stops is a bit more marked, although the variability among all speakers remains robust. A closer examination of individual informants suggests that, firstly, seven speakers favour the target affricate form in first place: five males (Francisco, Bernardo, Andres, Cesar, and Miguel) and two females (Cecilia and Melinda). Second, eight speakers favour the non-target stop over the affricate; of these three are males (Jamie, Sebastian, David) and five are females (Paola, Josefina, Consuelo, Caterina, and Carmen). So thus far, two tendencies have been determined: a significant number of speakers use the non-target stop in first place – most of them

females – and a similar amount favour the affricate – mostly males. The speaker Carola shows practically the same percentages for the target affricate and the fricativised, a pattern that is not replicated by any other speaker, with the possible exceptions of Miguel and Andres, who also show high rates of the fricativised.

Broadly speaking, the pattern of use of the three variants requires an analysis of other factors; sex, though, seems to be the most relevant factor.

4.2.3.3.1.2 Speakers' sex

In the variationist tradition, female speakers have shown different behaviour to male speakers. Valencia (1993-1994) showed that women favoured prestigious forms[120] over males and Cedergren (1973) showed that females used higher rates of the fricative variant.[121] The examination of this factor group revealed that similar patterns were found, and that the differences between males and females are statistically significant (*p-value = 0.040276; Chi-square = 4.21*); however, the intra-group behaviour of both sub-groups is dissimilar, hence this group is not analysed further.

4.2.3.3.1.3 Years of instruction in English

Hypothetically, increasing exposure to and experience with English would predict higher rates of use of canonical forms, as it was for the first variable examined. However, the intra-group analysis reveals that the behaviour of speakers is very different amongst each other, and, therefore they display non-homogenous behaviour. Hence, the analysis of this as a factor group cannot be considered as such, despite the fact that it was initially detected that fifth-year speakers showed the highest use of the target form (52% of the affricate) and that the differences were statistically significant.[122]

[120] The term prestigious refer to the Chilean prestige variant used by Valencia (1993-94) that, in my sample, can be identified as the stop.

[121] In Cedergren (1973), the fricative variant corresponds to the prestigious new form in the process of language change in Panama City speech.

[122] P-value = 2.723156×10^{-15}, Chi-square = 74.36

4.2.3.3.1.4 Socioeconomic group (SEG)

As mentioned earlier, social class has shown to be an important factor in the use of particular variants in Panama Spanish speakers (Cedergren, 1973). The social indexation of socioeconomic groups of the informants of my sample generates three groups: C2 corresponding to upper-middle class, whereas C3 and D correspond to middle-middle class and lower-middle class, respectively. A distinction amongst the three groups is expected; particularly, the two variants have shown different sociolinguistic values in Chilean Spanish, so a transfer of these patterns is predicted, i.e., lower-middle informants would use the "new" prestigious variant (stop) more than the other SEGs.

Table 4.2.4 Distribution of voiceless alveolar affricate (tʃ), per speakers' socioeconomic group, non-categorical speakers, groups C3 (middle-middle) and D (lower-middle) only (N=524)

	[ʧ] Affricate		[tʲ] Stop		[ˢʃ] Fricativised	
Social class	%	N	%	N	%	N
Middle-middle	50	148	24	71	26	76
Lower-middle	42	95	53	121	6	13
Total	**46**	**243**	**37**	**192**	**18**	**89**

From the sample, only C3 and D can be compared, as group C2 is made up of only two speakers (Miguel and David) who have very different linguistic behaviour, and, therefore, cannot be grouped.

The difference between middle-middle and lower-middle class is significant (*p-value = 3.721448 x 10^{-14}; Chi-square = 61.84*). The analysis reveals that most middle-middle-class speakers favour the target form over the non-target forms, whereas most speakers from lower-middle class favour the stop, that is, intra-group variation can be considered homogenous.

Table 4.2.4 above shows that the highest rates of the stop variant are in the lower-middle group (53%). Recall that the variant stop corresponds to the Chilean prestigious variant, which has not been reported (to this author's knowledge) in English; the results shown here suggest that speakers are hypercorrecting the use of (ʧ) by, instead of using the standard affricate, transferring the Chilean prestigious form to English.

The occurrence of hypercorrection in both L1s – English and Spanish – has been linked to lower-middle-class speakers (see e.g. Labov, 1966), which is also confirmed in this study for L2 situations. In this case, hypercorrection is explained by the use of the prestigious form over the target form, which is also target in the L1.

4.2.3.3.1.6 Summary of findings – social factors

From the analysis of the social constraints considered, and after the removal of the two categorical and exceptional speakers – Juan and Marcela-, the results suggest that:

- Voiceless alveolar affricate shows a considerable distribution across the three most relevant variants:[123] target affricate 46%, non-target stop 37%, and fricativised 18%.
- The hypothesis that most speakers would favour the canonical variant is confirmed in the overall results: the preferred variant is the affricate – standard in both English and Spanish – followed by the stop, which corresponds to the most prestigious of the three forms in Chilean Spanish. The hierarchy of use, therefore, is affricate > stop > fricativised.
- Individual speakers: the behaviour of the remaining speakers of the sample (two categorical speakers removed) is heterogeneous; seven informants favour the affricate, against eight speakers who favour the stop. One speaker favours similarly affricates (42%) and fricativised (41%) with one instance of difference (N=53 and 52, respectively).
- Speakers' sex: although gendered differences are statistically significant, intra-group variability showed that the distribution of the three variants for both males and females is not homogenous, preventing me from contrasting these as groups.
- Years of instruction in English: the most relevant finding in relation to this factor is that fifth years use the target form over the other sub-groups, as expected. Like with sex, the intra-group variability is not homogeneous; therefore, despite being significant, the examination of this factor is not valid in terms of comparative behaviour.

[123] The variant fricative was removed from the analysis due to the small amount of tokens.

- SEG: only groups C3 and D are examined; the most relevant fact is that the use of the non-target form shows comparable higher rates of use in lower-middle-class speakers than in middle-middle-class individuals, confirming, in relation to external factors, the hypercorrection hypotheses. The results for SEG are statistically significant.

However relevant these findings are, the examination of linguistic constraints is necessary to test for the formulated hypotheses of my study.

4.2.3.3.2 Linguistic constraints

The effect of internal linguistic constraints has already been justified for the first variable. Given the common grounds in literature and the similar characteristics of (ʃ) and (ʧ), four internal factor groups are considered: preceding and following segment, position within the word, and lexical item. Based on the results from previous studies, the predictions for these specific factors are: (1) a relevant effect of preceding segment, particularly of vowels with weaker forms, and (2) initial prosodic position, favouring strengthening.

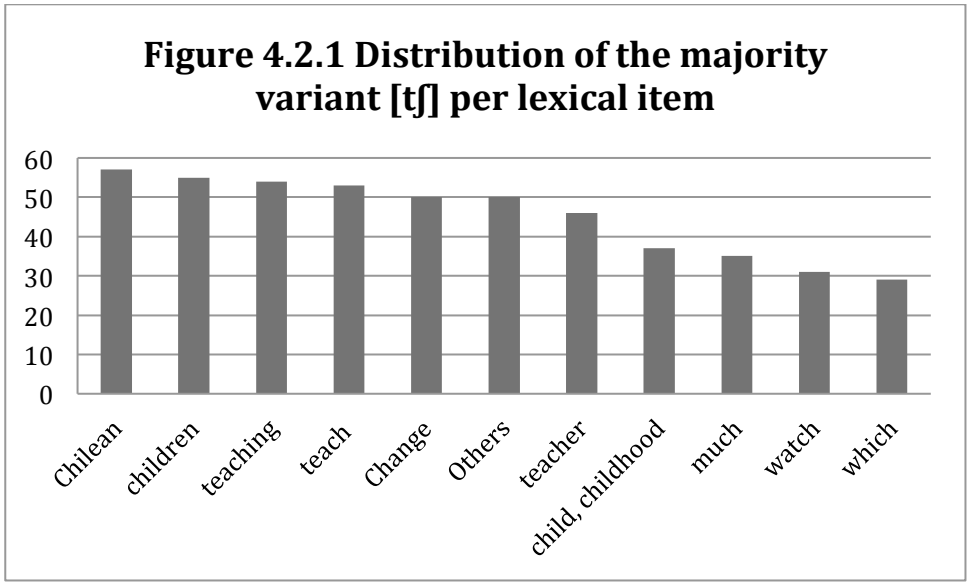
4.2.3.3.2.1 Lexical item

The analysis of individual lexical items is justified on the basis that some words share structural characteristics that may contribute to the similitude in use of a given variant (cf. Section 4.1.3.3.2.1, for the analysis of lexical item on voiceless alveolar fricative); for instance, we might expect that words such as *Chilean* and *children* show similar patterns of variation as both present the variable in initial position, the following segment is the same, and they are both open class words. In this section I will examine the use of the variable (ʧ) according to the lexical item in which it appears. Table 4.2.5 shows the distribution of (ʧ) for all lexical items considered (>15 instances) (N=618). The difference among the elements of this group is statistically significant (*p-value = 3.343941 x* 10^{-8}*; Chi-square = 74.47*).

Table 4.2.5 Distribution of voiceless alveolar affricate (tʃ) per individual lexical item (N=618)

	[ʧ] Affricate		[ʈʃ] Stop		[ᵗʃ] Fricativised	
Lexical Items	%	N	%	N	%	N
Chilean	57	20	20	7	23	8
Children	55	11	20	4	25	5
Teaching	54	26	35	17	10	5
Teach	53	24	31	14	16	7
Change	50	8	50	8	0	0
Others	50	73	39	57	10	15
Teacher	46	62	21	28	33	45
Child, Childhood	37	10	63	17	0	0
Much	35	30	48	42	17	15
Watch	31	9	52	15	17	5
Which	29	9	61	19	10	3
Total	**46**	**282**	**37**	**228**	**18**	**108**

Under the category *others* are included all those lexical items with less than 15 tokens. The words *teach, teacher and teaching* were separated in different categories as they present different following phonetic environments. The Figure 4.2.1 below displays the distribution of voiceless alveolar affricate in all lexical items with >15 occurrences; the figure displays in descending order the use of the affricate variant.



The main findings for the analysis of individual lexical items are:

- There are no categorical words, that is, all of them show some degree of variation.

- As shown in Figure 4.2.1 above, the highest percentages of the affricate variant (50% and above) occur in the lexical items *Chilean, children, teaching, teach* and *change* (plus the group *others*). All of these are content words and the variable occurs in three different positions: initial for *Chilean*, *children* and *change*, middle for *teaching* and final for *teach*; the last two lexical items share the same lexeme.
- The lowest percentages (<35%) occur in the words *much*, *watch*, and *which*. In all these three, the variant is found in final position, and two of them are function words.
- The analysis of the group *others*, reveals practically the same hierarchy shown by the overall distribution of the variable, i.e., the affricates are the preferred variant, followed by the stops and then the fricativised.

If content versus function words are analysed closely, they display the following behaviour:

Table 4.2.6 Distribution of voiceless alveolar affricate (tʃ) per word class (function and content)

	[ʧ] Affricate		[tʲ] Stop		[tˢʃ] Fricativised	
Word class	%	N	%	N	%	N
Content	**49**	237	33	163	18	89
Function	35	45	**50**	65	15	19
Total	**46**	**282**	**37**	**228**	**18**	**108**

Whereas function words, such as *which* and *much* seem to favour the stop, the affricate is used in higher rates in content words; this difference tested to be statistically significant (*p-value=0.001598253; Chi-square = 12.88*).

As differences in the articulation of function and content words have been found to be correlated to the position of the variable in the lexical item and other factors (see e.g. Bell *et al*., 2003; they examine frequency, predictability and utterance position), the other factors considered in this dissertation are examined in order to find the source of variation. I follow with the analysis of phonetic environment; as mentioned in the previous variable (Section 4.1.3.3.2.2), this factor is bound to lexical item, therefore it will be examined first individually and then in interaction.

4.2.3.3.2.2 Phonetic environment

4.2.3.3.2.2.1 Preceding phonetic environment

As a group, preceding phonetic environment is shown to be a powerful contributing factor in the variation of voiceless alveolar affricate (Cedergren 1973); also, with the previous variable examined, an effect of vowels on the use of the affricate variant was noted. An initial examination of all preceding segments found in the data set revealed that the patterns attested for the previous variable were not found here: there seems to be no pattern in terms of segments with similar articulation, i.e., vowels and consonants showed very dissimilar intra-group behaviour; furthermore, this dissimilar behaviour was even detected in minor categories, as in, e.g., plosives, fricatives, etc. Therefore, this factor group is discarded.

4.2.3.3.2.2.2 Following phonetic environment

Following a similar procedure as for the previous factor, I conducted a preliminary examination of all following segments found. This examination revealed that the intra-group behaviour was consistent, allowing the categorisation of the elements into larger categories. The table below shows the distribution of the three variants examined for vowels and consonants for following phonetic environment.

Table 4.2.7 Distribution of voiceless alveolar affricate (tʃ) in following phonetic environment: Pause, consonants and vowels (N=618)

	[ʧ] Affricate		[tʲ] Stop		[ˢʃ] Fricativised	
FPhE	%	N	%	N	%	N
Vowels	51	245	33	159	17	80
Consonants	30	34	49	55	21	24
Pause	14	3	67	14	19	4
Total	**46**	**282**	**37**	**228**	**18**	**108**

The difference among the three categories is significant (*p-value=0.00003510453; Chi-square = 25.78*). Again, there is an important amount of tokens in which the variable is followed by vowels (N=484; 78% of the total tokens). In following segment, there seems to be an important difference in the distribution of (ʧ): the affricate variant is

used more with vowels (51%, N=245) and the stop in pause contexts; however, the pauses represent only a small percentage of the total tokens.

Cedergren (1973) suggests that – in relation to the lenition of (ch) – vowels favour the lenition and consonants restrain it. This is consistent with my data, if we apply the same principle of weakening for the two variants: the stop is articulatorily stronger than the affricate, despite the fact that both present the same constituents: a plosive and a fricative.[124] This means that consonants contribute to the use of a strengthened variant, i.e., the stop, as opposed to vowels, which favour the use of the weaker form, i.e., the affricate.

In sum, only following segment appears to explain the variation of voiceless alveolar affricate, but as a factor, it is not robust enough to justify the variation by itself; therefore, it is necessary to examine interaction with lexical item.

4.2.3.3.2.2.3 Interaction of following segment and lexical item (word class)

Lexical items are examined in terms of class, in interaction with following segment. The cross-tabulation of these two factor groups is shown in Table 4.2.8.

[124] Recall that both are affricates, but for practical purposes I identify the standard form as affricate, and the non-standard as stop; refer to Section 4.2.3.1 Variants.

Table 4.2.8 Cross tabulation of following segment (vowels, consonants and pause) and word class (function vs. content words)

			Following phonetic environment							
			Vowels		Consonants		Pause		∑	
			N	%	N	%	N	%	N	%
Word class	Function	Affricate	15	23	29	56	1	8	45	35
		Stop	36	55	20	38	9	75	65	50
		Fricativised	14	22	3	6	2	17	19	15
		∑	65		52		12		129	
	Content	Affricate	19	40	216	50	2	22	237	48
		Stop	19	40	139	32	5	56	163	33
		Fricativised	10	21	77	18	2	22	89	18
		∑	48		432		9		489	
	∑	Affricate	34	30	245	51	3	14	282	46
		Stop	55	49	159	33	14	67	228	37
		Fricativised	24	21	80	17	4	19	108	17
		∑	113		484		21		618	

The cross-tabulation of following segment and word classes (function and content) shows that with function words, the use of vowels seems to favour the use of stops over affricates, whereas content words present a consistent use of affricates in both vowels and consonants; the use of stops, however, in content words is slightly higher in vowels and pause, despite the small number of occurrences (in fact, for content words, the use of affricates and stops is the same in following vocalic context). Thus, the variability exhibited by function in opposition to content words seems to be correlated to phonetic context. I will test again for the interaction of multiple factors in the Multivariate Analysis, Section 4.2.3.3.3.

During the examination of individual lexical items, it was noted that some of these words exhibited similar behaviour in relation to the position of the variable, in which final position showed higher rates of the use of stops and initial of the affricates; hence, I now turn to test for the interaction of position of the variable and lexical item.

4.2.3.3.2.3 Position within the word for individual lexical items.

The position of the variable within the word tested is statistically significant (*p-value = 0.0002925193; Chi-square = 21.17*). The distribution of position, as show in Table 4.2.9 below, indicates that both initial and middle positions favour the affricate (52% and 46%, respectively), whereas final position favours the stop (45%).

Table 4.2.9 Distribution of voiceless alveolar affricate (tʃ) in initial, middle and final positions (N=618)

	[ʧ] Affricate		[t] Stop		[tʃ] Fricativised	
Position	%	N	%	N	%	N
Initial	52	94	36	64	13	23
Middle	46	95	29	59	25	51
Final	40	93	45	105	15	34
Total	**46**	**282**	**37**	**228**	**18**	**108**

However, intra-group examination revealed that some individual lexical items show different behaviour to the group tendency (see Table 4.2.5 above), especially in final position. A cross-tabulation of these lexical items and the position of the variable is presented next to disentangle the variability of these two factor groups; the table is arranged showing first the lexical items in which the variable is in initial, then in middle, and then in final positions, as it has been argued that certain words that share structural characteristics present similar patterns of use (cf. analysis of voiceless alveolar fricative Section 4.2.3.3.2.1).

Table 4.2.10 Cross tabulation of position of the variable in the word and individual lexical items

Position of the variable in the word	Individual lexical items			N	%
	Initial	*Chilean*	Affricate	20	57
			Stop	7	20
			Fricativised	8	23
			∑	35	
		Children	Affricate	11	55
			Stop	4	20
			Fricativised	5	25
			∑	20	
		Others	Affricate	45	54
			Stop	28	34
			Fricativised	10	12
			∑	83	
		Change	Affricate	8	50
			Stop	8	50
			Fricativised	0	0
			∑	16	
		Child, childhood	Affricate	10	37
			Stop	17	63
			Fricativised	0	0
			∑	27	
	Middle	*Teaching*	Affricate	26	54
			Stop	17	35
			Fricativised	5	10
			∑	48	
		Teacher	Affricate	62	46
			Stop	28	21
			Fricativised	45	33
			∑	135	
		Others	Affricate	7	32
			Stop	14	64
			Fricativised	1	5
			∑	22	
	Final	*Teach*	Affricate	24	53
			Stop	14	31
			Fricativised	7	16
			∑	45	
		Others	Affricate	21	52
			Stop	15	38
			Fricativised	4	10
			∑	40	
		Much	Affricate	30	34
			Stop	42	48
			Fricativised	15	17
			∑	87	
		Watch	Affricate	9	31
			Stop	15	52
			Fricativised	5	17
			∑	29	
		Which	Affricate	9	29
			Stop	19	61
			Fricativised	3	10
			∑	31	

The examination of this cross-tabulation suggests that:

For initial position, whereas the words *Chilean* and *children* exhibit a higher use of affricates, *child (-hood),* does it of stops. The word *change* uses both fricatives and affricates equally, but the number of occurrences is very low (N=16 total).

In middle position, the variation is quite spread across the three variants, but the highest percentages of use occur in affricates.

In final position, the general tendency indicates that the variant used the most is the stop, with exception of the word *teach*, that uses similar percentages of the affricate as other words with the same lexeme (*teacher, teaching*). This suggests that position does not seem to justify the variability of voiceless alveolar affricate, but rather the different lexical items, i.e., structural similarities of words, seems to have a stronger effect. From this, it makes sense that the groups *teach, teacher, teaching* (same lexeme) and *Chilean, children* (same following segment, both content words) show a more similar behaviour than if we compare the results of the same lexical items classified in different factor groups, e.g. content vs. function words. Position, then, turned out to be a secondary factor, following individual lexical item; to test for the effect of the factors combined, a multivariate analysis is carried out below, in Section 4.2.3.3.3, but first a summary of the main findings drawn from the analysis of linguistic constraints is provided.

4.2.3.3.2.4 Summary of findings - linguistic constraints

The factor-by-factor analysis conducted above has revealed the following points about the effect of linguistic constraints in the use of (ʧ):

Lexical item

In relation to the analysis of individual lexical items, this group did not present categorical items, with all of them showing some degree of variation in two or three of the examined variants. The overall distribution suggests that affricates are the most used variant, followed by stops; the use of the fricativised variant occurs in only 18% of the tokens. When word class was examined, there was a noted difference for function and content words, with the first group favouring the stops and the second

the affricates; this difference was shown to be non-significant. As a potential conditioning factor inside lexical item, I examined the effect of preceding and following segment individually, and then in interaction. The results suggest:

Preceding segment

Intra-group behaviour is heterogeneous, which does not allow the examination of the elements of the group.

Following segment

The effect of following segment seems to be consistent, at least in the case of vowels for affricates. This group tested to be non-significant, but the results are consistent with previous studies. The interaction of following segment and lexical item reveals that, to some extent, word class is correlated to phonetic context.

Position

The position of the variable within the word was examined in interaction with the individual lexical items, as its distribution indicated that there was a substantial difference between initial, middle, and final positions; these differences are statistically significant. The results suggest that only middle position was consistent in the use of affricates over other variants (50% of use), which does not justify the variation of voiceless alveolar affricate on its own.

From the considered factors, the most robust seems to be the lexical items by themselves, particularly in reference to its structural characteristics, that is, how the similarities between some words – such as phonetic context, class, position, etc. – as a whole, have stronger effects on the variation than all these factors studied individually. In other words, as the source of variation does not seem to be robust enough for most factors, if considered individually, the examination of their interaction by means of multivariate analysis is required.

4.2.3.3.3 Multivariate analysis

Having completed the examination of all factors individually, and the exploration of their interaction by means of the cross tabulation of those factors that appear to be relevant to the variation of voiceless alveolar affricate, I present in this section the analysis of all factor groups combined, in order to uncover whether or not the interaction amongst the groups accounts for the use of the existing variants.

Before comparing the multivariate analysis of both variants, a separate multivariate analysis was completed to determine the variable rules for affricates and stops independently. This standard procedure (see e.g., Bayley, 1996; Tagliamonte, 2006), allows the researcher to establish the similarities between the variable rules, the order in which the several factors constrain the use of the variants, and in which range.

For each variant I consider all the factors examined previously; these are displayed in Table 4.2.10 and arranged from the highest to the lowest factor weight per factor group. Table below shows the independent multivariate analysis carried out only for the majority variant: affricate.

4.2.10 Multivariate analysis for voiceless alveolar affricate (ʧ), for the majority variant [ʧ]			
	Factor Weight (FW)	%	Total N
Input .46		46	618
Following segment			
Vowel	.56	51	484
Consonant	.34	30	113
Pause	.17	14	21
Range	**39**		
Preceding segment			
Vowel	[.56]	45	470
Consonant	[.34]	48	141
Pause	[.17]	57	7
Socioeconomic group			
Middle-middle class	[.55]	50	295
Lower-middle class	[.46]	42	229
Upper-middle class	[.46]	42	94
Lexical Item			
Much	[.39]	35	87
Teaching	[.59]	54	48
Teach	[.58]	53	45
Change	[.55]	50	16
Others	[.55]	50	145
Choose	[.50]	46	135
Chilean	[.62]	57	35
Children	[.60]	55	20
Child, childhood	[.41]	37	27
Watching	[.35]	31	29
Which	[.33]	29	31
Position of the variable in the word			
Initial	[.56]	52	181
Middle	[.51]	46	205
Final	[.45]	40	232

Table 4.2.10 above indicates the contribution of the factors to the application of the rule. As the input ("the overall tendency of rule application", Tagliamonte, 2006: 156) determines, the use of affricates is favoured (.46). The only factor group that was found to be significant is shown in bold and those which are not statistically significant are in square brackets. The range indicates the importance of the factor to the overall contribution of the application of the rule, that is, the higher the range, the greater the contribution (for greater detail on variable rule analysis, refer to Tagliamonte, 2006; Preston, 1996, Sankoff and Cedergren, 1974). Thus, I will only examine the groups that were found significant in the conditioning of (ʧ).

Following segment: with a range of .39, this indicates that the only factor that favours the application of the rules is vowels (FW = 56), with consonants and pause disfavouring it (FW = .34 and .17, respectively).

Other factor groups remain non-significant, which means that the variation of voiceless alveolar affricate is statistically conditioned by the linguistic factor *following segment*[125] only.

4.2.3.3.4 Summary and conclusions

The first finding in relation to the variation of voiceless alveolar affricate (ʧ) is the occurrence of four variants: an affricate [ʧ] (the standard form for both English and Spanish), a "stop" [tʃ] (a fudge form in which the emphasis occurs on the plosive constituent of the feature, prestigious in Chilean Spanish), a "fricativised" [tʃ] (also a fudge form, but with the emphasis on the fricative constituent) and a fricative [ʃ]. The hierarchy of use is affricate > stop > fricativised > fricative; the results confirm the hypothesis that the most used form would be the affricate.

In general, the results show that the predictions around the social factors found to constrain the variable selected are not particularly strong for sex, but significant, nevertheless, as both males and females use the target and the prestigious forms similarly.

For years of instruction in English, the only relevant finding is related to the highest use of the target form with fifth-year speakers, but the behaviour of the sub-groups is dissimilar, which prevents me from patterning the speakers' variation in terms of their progression on the acquisition of the target form.

The findings for socioeconomic group are interesting as they show that, as a group, the lower-middle-class shows higher rates of the prestigious form (the stop) over the target form, but intra-group analysis revealed that this apparent outcome was not homogenous for all speakers from the lower-middle class.

[125] Individual speakers, considered the fundamental piece of linguistic variation, cannot be considered a factor group as they cannot be *grouped*. What can be grouped are their characteristics that identify them in terms of sex, age, social class, etc., which are considered and examined as such. If individual speakers are discarded as a conditioning factor group, the only remaining group is following segment.

Regarding the linguistic factors, the effect of following segment was shown to be the most robust: vowels were the only element that favoured the use of the target form, which is, as predicted, consistent with the literature.

The other factor that seemed relevant was position, or at least the differences between the elements of the group were found significant – particularly for initial and final positions. The contribution of this group to the variation was not confirmed by the multivariate analysis as significant.

With regards to the hypotheses formulated for this variable, particularly the one related to the transfer of the Chilean prestigious variant to English, this was partially discarded, as, despite showing high rates of use, it is the target form which is the most used; therefore, the transfer does not occur. Furthermore, when examined from the interaction of all factors by means of the multivariate analysis, the variability of voiceless alveolar affricate appeared to be primarily motivated by linguistic factors, in particular by following segment.

I return to these results in the discussion in Chapter 5. I now turn to the analysis of the data for the third variable, voiced dental fricative (ð).

4.3 Voiced dental fricative (ð)

4.3.1 The variable

Voiced dental fricative (ð) is frequently represented in English orthographically by "*th*" and it occurs in a number of phonetic environments, as in *the, mother,* and *with*. This phoneme is considered rare in the world's languages (see e.g., Wells, 1982; Dubois and Horvath, 1998a; Maddieson, 2011). Unlike in English, there is no orthographic representation for (ð) in Spanish, nor has it been registered in the Spanish phonetic inventory. In Spanish, voiced dental fricative (ð) has been recorded as a variant for (d) in intervocalic and final positions (see e.g., Cepeda, 2001; Pérez, 2007); on the other hand, in English as L1 (e.g., Baker, 2008; Local, 1995) and as L2 (in SLA contexts – You, *et al*., 2005), one of the recurrent variants for (ð) is [d]. The examples below exemplify the cases for English as L1 (example 15), as L2 (examples 16 and 17) and for Spanish (example 18):

(15) *Them* /ðem/ has been documented, e.g., by Labov (1972: 78) as:

- an interdental fricative [ðem],
- an affricate [dðem], or
- a lenis dental stop [dem].

(16) *There*, its strong form /ðeəʳ/and weak form /ðəʳ/ (in Dutch learners of English, Wester *et al*., 2007) is documented as:

- [deəʳ] or [dəʳ], respectively (in initial position).

(17) *With* /wɪð/ can be produced as [wit] (in final position, in Dutch learners of English, Wester *et al*. 2007).

(18) *Cansado* (tired) / *amistad* (friendship) (my own examples, based on Pérez, 2007):

- standard/kan'sado/, /amis'tad/
- non-standard [kan'saðo], [amis'tað] or
- non-standard [kan'saØo], [amis'taØ]

This variable has been often found to be constrained by both social and linguistic factors; within the latter, particularly by phonetic environment, and by the interaction of word class (function versus content) and prosodic position (within the word and in turn). I now turn to the review of previous studies conducted on voiced dental fricative and their findings for English, Spanish and L2 situations.

4.3.2 Previous studies and specific hypotheses

4.3.2.1 Variation in English

In articulatory terms, the variation for (ð) in English has been documented by authors such as Strevens (1985) and Ladefoged (1996, 2001), who show that in the teaching

of accents and dialects such as the "received pronunciation" (RP)[126] of British English and American English (th)[127] /ð/ is described as "typically made with the tip of the tongue behind the upper front teeth" and as "interdental", respectively. Roach (2004: 241) describes /ð/ as "often a weak dental plosive with no detectable friction noise; the sequence /nð/ often assimilates to [n̪n̪]" (see examples below). Even though Roach's description belongs to RP, which is spoken by around 2% of the native English speaking population of the south-east of England, it shows that assimilatin to a different phoneme is possible, i.e., there is variation. Consistent with this, Baker (2008) – based on the studies of Ogden (1996), Manuel *et al.*, (1992), Manuel (1995), and Local (2003: 123) – states that "/ð/ can be produced as a dentally articulated nasal if there is a preceding nasal sound" as in example (19) below, "or a dentally articulated lateral if the preceding sound is a lateral" as in (20). Other studies on the (partial) assimilation of /ð/ to preceding consonant include: Shockey (1973, 1978) and Gimson (1989).

(19) "Oh, I have bad luck, because I'm finishing my *programme this* year"
(Carola, female, fifth year speaker).

(20) "… maybe it's too late in *school though*"
(Melinda, female, fifth year speaker).

Manuel (1995: 454) argues that there are three possible factors for the assimilation of /ð/:

- /ð/ is a "voiced non-strident fricative", where frication is "very weak in amplitude";
- /ð/ is a coronal, very "prone to assimilation or other weakening processes"; and
- /ð/ has a very limited distribution in English, mostly in function words, which are commonly unstressed.

[126]RP is the form (or accent) of British English taught to speakers of ESL/ EFL.

[127] Often referred to as (dh).

This means that the phoneme is easily assimilated and, therefore, more likely to vary especially in terms of its phonetic environment.

Catford (2001) describes the articulation for (ð) as [ð][128] (apico-dental fricative). According to him, the "approximants of other dento-alveolar types can be represented by using the ´opening´ diacritic,¨ thus /ð̪/ represents apico-dental approximant (2001: 88). He also states that this phoneme as a *wide fricative* is "bordering on the approximants", as in the Danish [ð̞] (2001: 125). Ladefoged (1996) also describes the articulation of /ð̪/ basically as laminal, but he states that dental fricatives can be both apical and laminal. Wolfram and Schilling-Estes (2006: 363) list a number of processes that frequently occur with *th*: it occurs as a stop at the beginning of the word (a phenomenon "spread across the full spectrum of vernacular varieties"), before nasals (in "Southern-based vernacular varieties, including Southern European-American and African-American vernacular varieties"), and in intervocalic position is produced as a fricative [v] or stop [d] (*brover, or broder* for *brother,* respectively).

In quantitative research, voiced dental fricative has been frequently linked to the phenomenon known as *th-stopping*, in which (ð) is commonly realised as a stop [d] (Wells, 1982). This feature has been described in different communities, with speakers of different social and ethnic backgrounds, and constrained by different social and linguistic constraints (for a complete account of this feature in the world Englishes see, e.g., Dubois and Horvath, 1998). For instance, it was found that the variation of voiced dental fricative is correlated to sex and network in Cajun speakers (Dubois and Horvath, 1998a, 1998b) and NSs of English (Labov, 2000); to ethnicity in London speakers of Cockney and non-Anglo (African, and Afro-Caribbean) background (Cheshire *et al.,* 2008) and in Chicago and New Jersey speakers of Italian origin (Eckert, 2008); and to dialectal identity for speakers from Buckie, Scotland (Smith and Durham, 2011, 2012). Labov examines the intersection of social factors (occupation, style, social class, age) in New York speakers of English as L1 and in connection to immigrant communities (1966, 2006), as does Wagner (2008) in Philadelphia with Irish and Italian descendants (age, social class, and ethnicity). Finally, the effect of social and linguistic factors has been tested by Milroy and Milroy

[128] In opposition to [ɵ], as in *mouth.*

(1978) in Irish speakers in Belfast (preceding and following segment, sex, and age), and by Mendoza-Denton (2007) in Latino immigrants in New York (phonetic environment, social grouping, code-switching).

The main findings from the above cited studies suggest that:

- In relation to class, *th-stopping* is generally stigmatised (Labov *et al.*, 1968), with the upper classes using more standard variants than the lower classes (Labov, 2006).
- Sex has been usually examined in interaction with other factors; for instance, in relation to network, females used vernacular forms more in closed networks than in open networks (Dubois and Horvath, 1998);
- In relation to ethnic background, Cheshire *et al.* (2004) suggest that males from ethnic backgrounds other than English are the highest users of [d] for /ð/ (especially in initial position, which is higher in Anglo than in non-Anglo speakers, and in females than in males – although in non-Anglo speakers *th-stopping* is reversed). Also, boys are identified as users of stops (in casual style – Eckert, 2008); and females are more likely to use the fricative over the plosives (Labov, 2006).
- Function words favour the use of the dental variant (Dubois and Horvath, 1998) and that *th-stopping* is higher in word initial position (Wagner, 2008).
- For word class (Manuel, 1995; Baker, 2008; Smith and Durham, 2011), most of the studies are restricted to function words in initial position (e.g., *they, then, them, the*); in my research I also include lexical words, such as *mother* and *brother,* in which the variable occurs in medial position.

Variation does not only occur across the different dialects of English, but may also take different forms depending on the ethnicity and group identity of speakers, socioeconomic class, network, style, articulation, phonetic environment and word class. From this literature, ethniticy is not relevant to my study of learners of English as a foreign language; however, it is interesting to observe how L2 speakers, mainly in immigrant settings, vary, as opposed to native speakers of English of different dialects.

The studies described above have determined that the variation of (ð) in NS of English is constrained by both internal and external linguistic factors; it is sought to test if this pattern is also reproduced in NNS.

4.3.2.2 Variation in Spanish

The variation of (d) in Spanish is relevant for this study for two reasons: first, it shows that systematic patterns of variation are found for this feature in the L1 of the speakers of the sample and, because of it, the use of [ð] in the L1 phonemic system may eventually interfere in the realisation of the English canonical [ð], and second, it offers a means of contrast to determine if the variation in L2 is constrained by the same factors in L1 situations.

Even though (ð) does not appear in the phonetic inventory of Spanish, Goldstein (2001: 55) and Martinez *et al.* (2003: 257) suggest, respectively, that this phoneme – or its approximant variant [ð̞] – occur in complementary distribution for /d/ in some dialects of Spanish, especially in intervocalic position, e.g. [deðo, aða].

Butragueño (2005) describes all allophonic realisations for /ð/ found in Spanish:[129] [δ] (voiced dental fricative), [ˠ] (weakened voiced dental fricative), [ʱ] (weakened voiceless-voiced dental fricative), [ð̞] (voiced dental approximant), and [ʶ] (weakened voiced dental approximant). From these descriptions it can be suggested that the phoneme /ð/ is part of the phonemic inventory of Spanish and it presents a considerable range of variation.

[129] Butragueño uses /δ/ for the IPA /ð/.

This variation is also present in Chilean Spanish, for which numerous authors have studied the linguistic behaviour of /d/. Pérez (2007) reports on the works of Oroz (1966) and Cepeda (1991), who found that /d/ is realised by means of: [d] in initial position and after nasals and laterals (*andar, saldo)*, apico-dental fricative [ð] in all other word positions (*enfadar)*, and elided in the endings –*ado, -edo,* and –*ido*, and in final position (*usted);* therefore, [ð] is an allophone of /d/. Wigdorski (1978, also reported by Pérez, 2007) argues about the tendency of the speakers of Santiago to aspirate or elide /d/ in intervocalic or final position (when followed by a consonant). This would have been relevant for my sample as a few of my interviewees are from Santiago and they might present some degree of variation different from the other local speakers; however, as there have been no studies carried out on this area of Chillan, we cannot predict any particular linguistic behaviour for local speakers.

Pérez (2007) also reports on the works carried out by Cepeda (1994), Cepeda and Poblete (1993) and Poblete (1995) in relation to the group /b-d-g/. According to Pérez, this group presents a general weakening in intervocalic position, and /b/ and /d/ are elided more frequently in preceding stressed vowel and following unstressed vowel, in suffixes (bound morphemes) more than in free morphemes, and in male, younger, lower-class speakers.

In a phonetic-stylistic study of the group /b-d-g/ in the speech of Chilean news programmes, Pérez (2007) found that there is considerable stylistic variation in the realisation of the three variables. In relation to (d), three variants were identified: closed approximant (in this case, the voiced dental fricative), approximant, and zero realization (elision). The overall distribution of the three variants is as follows:

Table 4.3.1 Overall Distribution, Phonetic-Stylistic study of plosives in Chilean News (Pérez, 2007)

Closed Approximant	28,8%
Approximant	47,0%
Elision	24,1%

In terms of stylistic variation, two styles were recognised: non-spontaneous speech (the actual reading of the news and the contextualisation of images or videos being broadcasted) and spontaneous conversation (remote or in-site interviews). The distribution of the variants according to style is shown in Table 4.3.2:

Table 4.3.2 Distribution of the variants according to style; phonetic-stylistic study of plosives in Chilean News (Pérez, 2007)		
	Style A "non-spontaneous speech"	Style B "spontaneous speech"
Closed approximant	30,1%	24.0%
Approximant	52.5%	26.8%
Elision	17.4%	49.2%

According to this data, plosives tend to be fricativised in intervocalic position, and also vary depending on the style of the speech. These results indicate that it is possible, at least in intervocalic position, to find the modification of our variable under study.

4.3.2.3 Production (and variation) in L2 settings[130]

In a study on the linguistic performance of Italian immigrants in Canada (not from the variationist methodology perspective, however), Flege, Munro and MacKay (1995d) examined (ð) in word initial position. According to Flege *et al.* (1995d: 6) (ð) is not present in the Italian phonetic inventory; in the study, the production of (ð) varied between native English speakers and native Italian speakers who had been learning English for nine to twenty-one years. The results showed that "incorrect realisations of (ð) were usually heard as [d]" Flege, Munro and MacKay (1995d: 7); for the authors this means that native Italian speakers failed to recognise the target phoneme, which was replaced by [d]. Flege et al. (1995) describe the following as possible causes for these non-standard realisations: the habit formation in relation to the acquisition of a second language, the loss of motoric ability (i.e., articulatory ability) to produce certain consonants, the failure to perceive the phonetic differences, or possible phonetic system effects (such as difficulties keeping two systems separate). Another reason for the replacement of this feature by a close sound is that, as Maddieson (2011) suggests, dental and alveolar forms such as (ð) occur in 43 (or 7.6%) of the world´s natural languages, but it has a worldwide distribution, making it a rare

[130] Most of the studies quoted in this section are not variationist, since they do not use the quantitative paradigm described by Labov (1966, 1972); they study the production of particular features in speakers of English as L2, from different L1 backgrounds, mainly at acoustic production level.

feature. Maddieson (2011) also states that in Spanish, the grapheme "d" is usually represented by (ð), except in initial position and next to certain consonants.

Similar results to Flege *et al.* (1995d) are shown in You *et al.* (2005), who analyse the production of consonants in children with Spanish-accented English. They propose as possible pronunciation for the variable (dh) the phoneme [d], based on articulatory similarities between these two sounds in English and Spanish. Their results determined that 34.6% of the substitutions of /ð/ occurred with the phoneme /d/,[131] with the possible exception of the allophone /dh/[132] in final position; this last phenomenon is also present in Spanish.

Wester *et al.* (2007) suggest that the substitution of dental fricatives in Dutch learners of English occurs in initial position largely with /d/ and in final position with /t/ (e.g., *with,*[133] *seethe, bathe, teethe*). This substitution is explained by the paradigmatic relationship between the "(non-sibilant) dental fricatives and their dental oclusive correlates" (Gamkrelidze, 1975), which states that non-existent phonemes in a phonetic inventory are replaced by their closest phoneme, in terms of manner, place of articulation or voicing. Gamkrelidze (1975) argues that in most languages there is a sort of equilibrium of phonemes, and if in a given language there is a pair /p/ and /b/ and one of these phonemes is absent (e.g., /b/) this will occur as an allophone of the other phoneme (e.g., /p/). If we compare the IPA charts for English[134] and Spanish[135] (Tables 4.3.3 and 4.3.4 correspondingly), we can identify the "lack" of correspondence for a dental fricative,[136] which can support our assumption. Based on Gamkrelidze's claim, I may hypothesise that if students fail to recognise/use the fricative variant, it is more likely that they will use the plosive, not

[131] This study is carried out in relation to the modelling of non-native speech to improve the performance of automatic speech recognition systems, and not from the variationist perspective, and it is based on the comparison of Spanish and English phonetic symbols. They attribute the differences in pronunciation to "acoustic phonetic level transfer" and to "orthographic knowledge transfer" (You *et al.* 2005).

[132] Their notation.

[133] In Standard English.

[134] ©IPA (revised 2005).

[135] Martínez-Celdrán, Fernández-Planas and Carrera-Sabaté (2003).

[136] Although this chart is for Castilian Spanish, it is mostly applied to all Spanish dialects. The possible exception is the phoneme /ϴ/, as its use is mostly peninsular.

because of the manner of articulation, but as its counterpart in the place of articulation.

Table 4.3.3 Phonetic chart -English consonants (RP) [137]								
	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b			t d			k g	
Affricate					tʃ dʒ			
Fricative		f v	θ ð	s z	ʃ ʒ		(x)	h
Nasal	m			n			ŋ	
Lateral				l				
Approximant	w				r	j		

Table 4.3.4. Phonetic chart - Spanish consonants						
	Bilabial	Labio-dental	Dental	Alveolar	Palatal	Velar
Plosive	p b		t d			k g
Affricate					tʃ ʄ͡ʝ	
Fricative		f	θ	s	ɲ	x
Nasal	m			n		
Lateral approximant				l		
Trill				r		
Tap or flap				ɾ	ʎ	

As can be seen from these studies, the production of L2 features has been shown to be related to the presence/absence of the phoneme in the L1 phonetic inventory of the interviewees. We can hypothesise, then, that in my sample the status of EFL of the speakers can be correlated alongside the linguistic factors in the variation of (ð), that is, the speakers would transfer their use of L1 features when speaking English.

[137] Retrieved from https://www.llas.ac.uk/materialsbank/mb081/page_07.htm

4.3.2.4 Hypotheses

Based on the literature reviewed for English and Spanish, along with the reported use of the feature in NNS of English, it is expected that the systematic variation of voiced dental fricative (ð) exhibited in NS of English is also found in this specific group of NNS. In particular, it is expected that (ð) is constrained exclusively by the linguistic environment, especially in intervocalic position or when it is preceded by nasals and laterals.

However, this variable has been shown to be constrained by linguistic internal factors only; if the variation in NS of Spanish is transferred to English, it might also be expected that an important contribution would be years of instruction in English (towards the progressive acquisition and use of the canonical English form), as well as social class.

4.3.2.5 Exclusions

During the coding process, it was noted that some instances of the studied variable occurred in neutralisation contexts. These are defined as the occurrence of "tokens in which independent processes exist which make the reliable identification of the variant under investigation difficult (or near impossible)" (Tagliamonte, 2006: 91). This refers to those contexts in which the variant is preceded or followed by a similar (or equal) phoneme or variant (in the case of phonetic variation), as in example (21) below:

(21) "My relationship with, *with* [th6][138] teachers was good, in general"

(Bernardo, fifth-year male).

In (21) the informant uses the laminal – dentoalveolar plosive variant (see variants Section 4.3.3.1), occurring at the end of *with*, and is followed by the word *teacher*. Here, it is very difficult to discriminate the variant from the following segment, therefore it cannot be included as a context. Other studies have excluded all realisations other than fricatives or plosives (Smith and Durham, 2011), which is also

[138] See Section 4.3.3.1 for a description of the articulatory characteristics of each variable.

the case in my study; if the manipulation of the data requires it, further exclusions will be made.

4.3.3 Variants and distribution

4.3.3.1 Variants

In my research, voiced dental fricative (ð) shows nine variants – as shown below with their corresponding codes and examples:

- (ð) deletion: also referred to as the zero realization of (ð); it is coded as [th0], as in "*you love them, so you cook for* them *[th0], but it's love!"* (Paola, first-year female).
- Laminal-interdental fricative, coded as [th1]: "I turn off the light and always I see them [th1], photograph- photos" (Cristina, first-year female).
- Laminal – dentoalveolar fricative, identified in the data as [th2]: "the *[th2] thing's I don't read"* (Caterina, third-year female).
- Apical – interdental fricative, this will be referred to as [th3]: "*how do you say* the *[th3] person who help* [sic] *to other people? (*Marcela, first-year female).
- Apical – dentoalveolar fricative, coded as [th4]: "the only way I could have a little bit of conversation with [th4] him was using English" (Andres, fifth-year male).
- Laminal – interdental plosive; this variant is identified as [th5]: "then [th5] I said why is everybody looking at me? Oh, my God, everybody is looking!" (Carola, fifth-year female).
- Laminal – dentoalveolar plosive, referred to as [th6]: "it's a very good movie and this [th6] is an epic movie" (Bernardo, third-year male).
- Apical – dentoalveolar plosive, coded as [th7]: "they [th7] have seminars and practise their English." (Carmen, fifth-year female).
- Voiceless laminal – dentoalveolar fricative, identified as [th9] in the coding: "(…) and I could communicate with [th9] her in English, and I felt good about it" (Melinda, fifth-year female).

The distribution of these nine variants – and how these are manipulated (in methodological terms) and analysed – are presented in the section below.

4.3.3.2 Distribution

As with the previous variables, in this section I will describe the frequency of use of the different variants for the variable (ð), according to (1) social factors – individual speakers, sex, SEG, and years of instruction in English – and (2) linguistic factors – preceding and following phonetic environment, position of the variable within the word, position in turn, and lexical item (open- versus closed-class words).

4.3.3.2.1 Distributional data description for /ð/

4.3.3.2.2.1 Overall distribution

Nine variants were initially identified. This is, in itself, an important finding, as it evidences that, as described in the literature, there is an important range of variation in non-native speakers of English, just as there is in native speakers – as shown by Labov (1966) in New York, Dubois and Horvath (1998) for Cajun English, Timmins, Tweedie and Stuart-Smith (2004) for Scottish English, Wester *et al.* (2007) in English by Dutch L1 speakers (see Section 4.3.2), and Smith and Durham (2011) for Doric speakers of English. The variation has been tested to be constrained by linguistic and social factors.

The identification of the variants – introduced in Section 4.3.3.1 – was made by means of an auditory analysis, as in many variationist studies (see, e.g., Milroy, Milroy, and Docherty, 1999). All possible articulations were first identified from the literature and then the researcher reproduced the exact sound produced by the informants, identifying the closure of the variant (fricative, plosive), the place of articulation (interdental or dentoalveolar) and, finally, whether the obstruction was made by the blade or the tip of the tongue (laminal or apical).[139] The distribution of the nine variants for (ð) is displayed below:

[139] This methodology is similar to the one used by Abramson and Lisker (1972) in a recognition experiment applied to Spanish-English bilingual speakers. In this, the individuals were asked to identify from a recording a set of stops and match them to their orthographic representations. See also Lawson, Scobbie and Stuart-Smith (2011).

Table 4.3.5 Overall distribution for voiced dental fricative (ð), all variants considered, all speakers (N= 1353)

	N	%
[th0] (ð) deletion	1	0.1
[th1] Laminal- interdental fricative	819	60.5
[th2] Laminal – dentoalveolar fricative	76	5.6
[th3] Apical – interdental fricative	4	0.3
[th4] Apical – dentoalveolar fricative	55	4.1
[th5] Laminal – interdental plosive	90	6.7
[th6] Laminal – dentoalveolar plosive	210	15.5
[th7] Apical – dentoalveolar Plosive	88	6.5
[th9] Voiceless laminal – dentoalveolar Fricative	10	0.7
Total	**1353**	

In this table, we can see that the laminal-interdental fricative variant ([th1] - the English canonical variant for /ð/) accounts for almost 61% of the instances, followed by the laminal – dentoalveolar plosive variant ([th6] - the Spanish standard use for /d/), with almost 16% of the instances. As these comprise 76% of the total tokens, a relatively small proportion of use of the other seven variants was produced. Following the traditional method of collapsing small Ns into larger categories (a procedure usually followed for preceding and following phonetic environments; see e.g., Cedergren, 1973; Guy, 1988; Bayley, 1996; Major, 1996; Tagliamonte, 2006), I decided to collapse the variants into two groups: fricatives and plosives, with the exception of two variants, which will be hereafter excluded from the analysis: [th0] ((ð) deletion, N=1) as it does not fit in any of the two broader categories and it only occurs once, and [th9] (voiceless laminal – dentoalveolar fricative, N=10), as this variant is the only voiceless sound, which appears in a very restricted setting (i.e., final position, for the lexical item *with*). The decision to collapse the variants into two major groups is mainly justified in articulatory terms; more specifically, it is based on the manner of articulation (see also Smith and Durham, 2011). The distribution of voiced dental fricative is presented in Table 4.3.6.

Table 4.3.6 Distribution of voiced dental fricative (ð) into fricatives and plosives (N= 1341)

	Fricatives	Plosives	Total
N	953	388	1341
%	71	29	

When these groups are merged, the overall distribution (OD) shows that 71% of the occurrences are fricatives and 29% plosives, as featured in Table 4.3.6 above. Despite the fact that the category of fricatives merges different articulations for /ð/, I will

refer to this as the target form, as a large proportion of it incorporates the canonical /ð/, in opposition to the group plosives, which is non-target. Thus, the target form is the majority variant.

4.3.3.3 Factoy-by-factor analysis

The following section presents the results drawn from the analysis of the selected social and linguistic factors. At the end of each section, a summary of the main findings is included.

4.3.3.3.1 Social constraints

As with the two previous variables, all factors mentioned in Section 3.4.3.2 were included. From this preliminary analysis, two factor groups were discarded:

4.3.3.3.1.1 Sex

The analysis reveals that gender differences do not seem to constrain the variation of (th): the use of both fricatives and plosives is fairly similar for both sexes, as shown in Table 4.3.7.

Table 4.3.7 Distribution of voiced dental fricative (ð) in all variable speakers, arranged by sex (N=1341)

	Fricatives		Plosives	
Speakers	%	N	%	N
Females	76	538	24	169
Males	66	415	35	219
Total	**71**	**953**	**29**	**388**

With females using the target form 76% and males 66%, the difference between the two groups of speakers for this factor group tested to be statistically significant (*p-value: 0.00002346272, Chi-square = 17.89*). Also, the intra-group analysis revealed that the individual speakers exhibited dissimilar linguistic behaviour, therefore they could not be grouped.

4.3.3.3.1.2 Socioeconomic group

The differences between the three social groups do not seem to constrain the variation of voiced dental fricative, as shown in Table 4.3.8.

Table 4.3.8 Distribution of voiced dental fricative (ð) in all variable speakers, arranged by SEG (N=1341)

	Fricatives		Plosives	
Speakers	%	N	%	N
Middle-middle	74	485	26	169
Lower-middle	70	311	30	136
Upper-middle	65	157	35	83
Total	**71**	**953**	**29**	**388**

With middle-middle class using the target form 74%, lower-middle class 70%, and upper-middle class 65%, the differences for this factor group tested to be significant (*p-value=0.02663258, Chi-square = 7.25*). Intra-group analysis also revealed the heterogeneous behaviour of its speakers, therefore their inclusion as part of a group makes no sense.

All remaining factor groups, i.e., individual speakers and years of instruction in English, are examined in the following sections.

4.3.3.3.1.3 Individual speakers

As mentioned in the section on the factors considered (3.4.3.2.4), the study of individual speakers is justified as it helps understand the behaviour of the group, either as a mirror of it, or as an influence for the other members of the speech community (see, e.g., Labov, 2001; Stuart-Smith and Timmins, 2010). The distribution of voiced dental fricative is shown in Table 4.3.8 for all individual speakers only for fricative variants; no categorical speakers with more than 95% of use of one of the two variants was detected, which means that the variable presents a good distribution across the two variables; Table 4.3.9 is arranged from the highest to the lowest percentage of use of the target form. Speakers shown in bold belong to first year, those in italics belong to third year; and normal font is used for fifth-year speakers.

Table 4.3.9 Distribution of voiced dental fricative (ð) for fricatives, for all individual Speakers (Total N= 1341)		
	Fricatives	
Speaker	%	N
Melinda	93	74
Josefina	*92*	*78*
Francisco	**92**	**34**
Caterina	*84*	*52*
Juan	*83*	*70*
Paola	**78**	**57**
Carola	78	119
Bernardo	*76*	*64*
Carmen	73	54
Sebastian	**73**	**49**
Miguel	65	44
Cecilia	**64**	**47**
Andres	64	46
Consuelo	*63*	*37*
Cesar	62	45
David	*49*	*43*
Marcela	**42**	**20**
Jaime	**33**	**20**
Total	**71**	**953**

The first major finding is that fifteen out of the eighteen speakers from the sample use the target variant over 50% of the time. The three remaining speakers are from first year (Jaime and Marcela) and third year (David). The fifteen speakers that prefer the fricative group are well distributed in terms of years of instruction in English (and sex, but sex is no longer considered as a contributing factor, as justified in Section 4.3.3.3.1.1 above).

4.3.3.3.1.4 Years of instruction in English

In the previous variables examined, it was noted that the use of target forms was progressive together with the time of exposure of the individuals to English, but as a factor group, years of instruction in English has not been tested as the most powerful contributing factor to the overall variation of the sample. From Table 4.3.10 we learn that this is a quite heterogeneous group in terms of percentages of use, so grouping them could obscure the inter-group results. However, this, as a factor group, tested to be statistically significant (*p-value=0.0004426322; Chi-square = 15.45*). Table 4.3.10 shows the tendency of the three groups, from the highest to the lowest, to use the fricative.

Table 4.3.10 Distribution of voiced dental fricative (ð) for fricatives for all individual Speakers, per speakers' years of instruction in English (N= 1341)		
	Fricatives	
YIE	N	%
Third	344	75
Fifth	382	74
First	227	63
Total	**953**	**71**

When we compare first- and third-year speakers there seems to be a clear progression towards the use of fricative variant over plosives. This progression seems to stop in the third year. As shown, the percentages for third- and fifth-year speakers are the same. This could suggest that the feature might be fossilised in third year; however, it is very difficult to test for this particular prediction, given the nature of the data. Besides, as already mentioned, the intra-group examination suggests that there seems to be no direct relation between the use of a particular variant and the years of instruction in English in terms of frequency of use. Further analyses are, therefore, required to test for the interaction of factors, as the effect of years of instruction in English, however significant, does not seem to provide a strong explanation for the variation of voiced dento-alveolar fricative.

4.3.3.3.1.5 Summary of findings – social factors

The main finding for social constraints is linked to individual speakers. Two social factors (sex and socioeconomic group) did not seem to contribute to the variation of voiced dental fricative, therefore they were not examined (as justified in Section 4.3.3.3.1). In relation to the years of instruction in English, this group tested to be significant, but the intra-group analysis revealed that all subjects could not be grouped together as they exhibited different linguistic behaviour. I now turn to the examination of linguistic factors.

4.3.3.3.2 Linguistic constraints

4.3.3.3.2.1 Lexical item

To test whether the results obtained for preceding and following segment and for position in turn and in the word are justified in terms of the lexical items in which voiced dental fricative occurs, I examine all lexical items; word class was initially

considered, but a closer analysis of the lexical words revealed that all content words in the corpus[140] presented categorical, or near categorical, use of the fricative (91%), with the exception of the word *clothes*, which presented categorical use of the plosive. All lexical items that account for less than 30 instances were re-grouped in the category *other*s, as justified for the previous variables.[141] I decided to collapse all <30 words also to allow the completion of multivariate analysis, as empty cells are avoided. Table 4.3.11 below shows all >30 lexical items, for all speakers, including group *other*s (<30); the numbers are arranged from highest to lowest percentage of occurrence of the target form (fricative).

Table 4.3.11 Distribution of voiced dental fricative (ð) in fricatives, for all lexical items >30 (N= 1341)

	Fricatives	
Lexical item	%	N
Another	98	48
Other - otherwise	98	110
Others (<30)	86	156
Them	81	79
With	81	121
There - therefore	72	82
Their	71	22
Those	63	26
That	61	94
Than	60	24
The	57	96
This	55	42
They	42	53
Total	**71**	**953**

[140] Content words: *brother, mother, father, further, together, gather, weather, clothing, rhythm*, and *clothes*; N=107. All these lexical items occur less than 30 times, therefore they are included in the category *others*. Basically, all lexical items examined individually are function words.

[141] In the previous two variables I collapsed all lexical items with less than 15 tokens; in this variable, most of the words with less than 15 occurrences were categorical and near-categorical, therefore I decided to extend it to all categorical and near categorical lexical items < 30.

The first confirmed finding is that the use of the fricatives is preferred in almost every lexical item, with the exception of the word *they* (plosives: 58%, N=73). The two first items from the list, *another* and *other*, are categorical; in these two items, the variable is in middle position (also intervocalic position).

The remaining words present the variable in initial position, with the exception of *with*, which, in terms of percentages, shows the same behaviour as *them*, so therefore, position should not be conditioning its use (the analysis of position as a factor is presented in Section 4.3.3.3.2.3 below). All examined words are function words. Content words, given the small numbers of occurrence, were collapsed into the group *others* (see footnote 145), but as a group, they do not show a marked difference with function words, that is, both content and function words show similar use of fricatives.

To understand the use of the plosives, therefore the origin of variation, an examination of the other linguistic constraints is required.

4.3.3.3.2.2 Phonetic environment

The analysis of both preceding and following segment is presented in the following sections. Also, an analysis of the interaction of both contexts is incorporated to examine in particular the effect of intervocalic position.

4.3.3.3.2.2.1 Preceding phonetic environment

As with the previous section, all preceding and following phonetic environments were collapsed into three groups: consonants, vowels and pause. The following examples show these three contexts:

- Consonants

  "(…) the true <u>is that</u> I'm don't see very often movies."
  (Cecilia, first year speaker).

- Vowels

"(„„) my grandfather died about, let's see, six years ago."
(Carola, fifth year female speaker).

- Pause

"The thing's I don't read."
(Melinda, fifth year female speaker).

The collapsing of phonological environments has been justified for the two preceding variables (see e.g., Section 4.1.3.3.2.2.1); it follows the rationale that, first, it simplifies the analysis process, and, second, the literature suggests that the effect of consonants is different to vowels, with the second favouring processes such as deletion or weakening and consonants the strengthening of particular phonemes. For instance, Dubois and Horvath (1998) initially considered type of consonant or vowel, but they finally coded for the categories consonants, vowels and pause.[142] In other variables (e.g., (ʃ)) it has been shown that intervocalic position tends to favour the use of fricatives; this phenomenon has been described widely in phonetics (see Section 3.4.3.2.5), therefore, similar results are expected here: that vowels favour the use of fricatives, as consonants restrain it. The results for the analysis of the effect of preceding segment on the variation of (ð) are shown in Table 4.3.12.

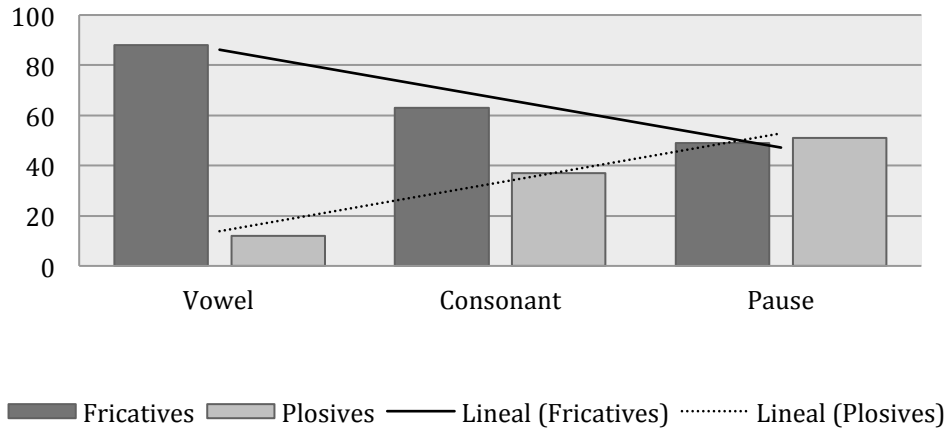
Table 4.3.12 Distribution of voiced dental fricative (ð) for fricatives in preceding phonetic environment: Vowels, consonants, and pause (N= 1431)

	Fricatives	
PphE	N	%
Vowel	505	88
Consonant	314	63
Pause	134	49
Total	**953**	**71**

The differences are statistically significant (*p-value = 9.178809 x 10^{-35}; Chi-square = 156.75*); vowels favour the use of fricatives (88%, N=505) over plosives (12%, N=69), as predicted. For consonants, the same hierarchy of use is exhibited, but the difference is not as marked (63%, N=314 for fricatives, versus 37%, N=181), as displayed in Figure 4.3.1 below.

[142] This factor group tested to be non-significant for the variation of (dh) in Cajun English.

Figure 4.3.1 Distribution of voiced dental fricative (ð) in preceding phonetic environment



These results suggest that, as in Pérez (2007), vowels favour the use of fricatives, which are followed by consonants with a less marked difference. Pause seems to have no effect on the use of fricatives (49%) or plosives (51%). This suggests that preceding segment is important in the realisation of the variable: with no preceding segment, the probabilities to produce a fricative or a plosive are similar. This last analysis (on preceding pause) is valid for the study of the position of the variable in turn, that is, when the variable occurs in turn initial (at the beginning of an utterance, for instance), the possibility of occurrence of a fricative is very much like that of a plosive (see this point below in Section 4.3.3.3.2.3). This will be tested for in the multivariate analysis. This factor, thus, needs to be examined in interaction with other factors, such as position in the word and in turn.

4.3.3.3.2.2.2 Following phonetic environment

As with preceding phonetic environment, following segment is also examined in categories, in this case, only vowels and consonants. Pause is excluded from the analysis as following pause presented only one token. The following examples illustrate the use of voiced dental fricative for vowels (22) and consonants (23):

(22) "(…) <u>they</u> have a reputation, and so <u>that</u> was important for me."

(Paola, first-year speaker)

(23) "(…) <u>with my</u> family, mainly, um, sometimes I meet my friends"

(Carola, fifth-year female)

As the effect of vowels has been recurrently found in variationist research as conditioning factor in the fricativisation of plosives and affricates (see e.g., Hickey, 1999; Cedergren, 1973), a similar effect to preceding segments is expected: fricatives being used more frequently when followed by vowels. In relation to the effect of consonants, Cedergren suggests that these restrain the use of fricatives, favouring the use of stronger articulation forms (such as plosives or affricates), thus, similar results are expected here. The findings for following segment are displayed in Table 4.3.13.

Table 4.3.13 Distribution of voiced dental fricative (ð) for fricatives, in following phonetic environment: Vowels and consonants (N= 1341)

	Fricatives	
FPhE	N	%
Consonant	91	80
Vowel	862	70
Total	**953**	

This group is statistically significant (*p-value=0.04056771; Chi-square = 4.19*); the effect of both consonants and vowels in following segment is relevant for the use of fricatives, which do not seem to be restrained by following consonants.

Table 4.3.14 Cross tabulation of preceding phonetic environment (vowels, consonant and pause) and following phonetic environment (vowels and consonants) (N= 1341)

			Preceding phonetic environment		
			Vowel %	Consonant %	Pause %
Following phonetic environment	Vowels	Fricative	90	63	49
		Plosive	10	37	51
	Consonants	Fricative	80	--	--
		Plosive	20	--	--

If examined in interaction, preceding and following segment confirm that the effect of intervocalic position in the use of fricatives is important, especially when compared to the effect of C_V (e.g., in the examples "can you *see there*?" versus "*like that"*). Similar results for the effect of following segment indicate that, as a constraining factor, preceding segment is not as strong as following segment, which needs to be tested by means of multivariate analysis; this is conducted at the end of the factor-by-

factor, in Section 4.3.3.3 below). Now I turn to the factor groups position in the word and in turn.

4.3.3.3.2.3 Prosodic position

Variationist research has shown that in initial position certain segments, such as fricatives, tend to be strengthened (see e.g. Wester *et al*., 2007) when compared with their occurrence in other positions. As shown in Section 4.3.3.3.2.2.1, the effect of a preceding pause (i.e., the zero occurrence of a preceding segment before a variable instance) showed that the likelihood of occurrence of a fricative is similar to that of a plosive in initial turn position. In this set of data, the results are similar to the ones shown by the literature: 46% (N=27) favour the fricatives vs. 54% (N = 32) for plosives.

A preliminary examination of the position in turn also showed that other positions favour the use of fricatives over plosives (72%, N=926).[143] These results indicate that the source of variation can, to some extent, be justified from the position of the variable within the word; thus, an examination of it is presented below.

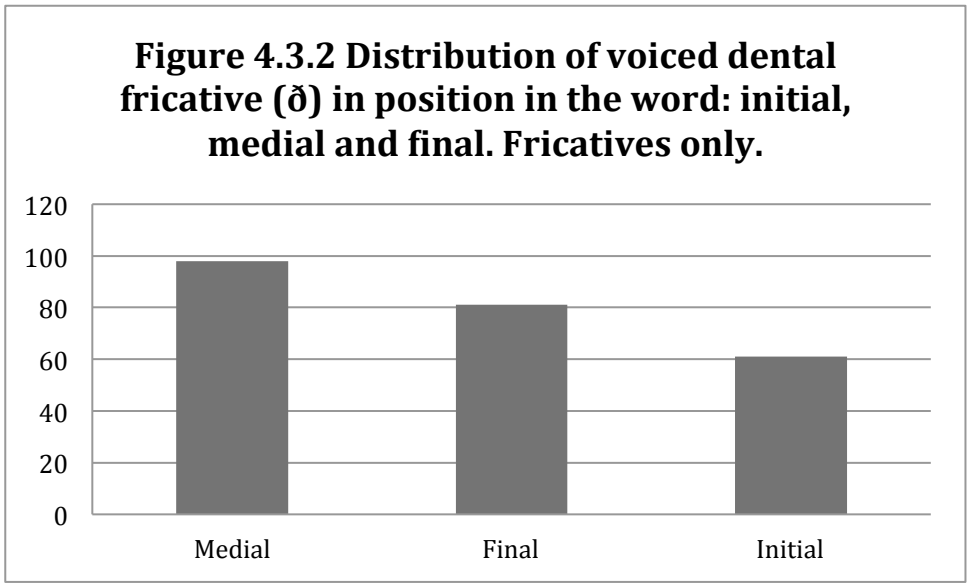
4.3.3.3.2.3.1 Position within the word

To corroborate if the use of a determined variant (or group of variants) can be linked to other factors than phonetic environment and to test whether the effect of position is in general confirmed, I analyse the position of the variable within the word, i.e., initial, middle and final positions. Studies (e.g., in articulation variation, Fougeron and Keating, 1997; Fougeron, 1999, 2001; Cho, 2004) have shown that the strongest position is initial, i.e., the segments occurring in initial position tend to be strengthened as opposed to those occurring in middle and final position, and are very much likely to suffer weakening processes such as lenition or deletion (see e.g., Cedergren, 1973; Wester *et al.,* 2007; and Pérez, 2007). In final position, only one lexical item was found: *with*. (The examination of the interaction of all lexical items and position is shown in Table 4.3.15).

[143] The results of the significance test tells us that position in turn is significant, with a *p-value = 0.00002265917, Chi-square = 17.95.*

Table 4.3.15 Distribution of voiced dental fricative (ð) for fricatives, in prosodic position in the word: initial, middle and final positions considered (Total N= 1341)		
	Fricatives	
Position	N	%
Medial	282	98
Final	121	81
Initial	550	61
Total	**953**	**71**

Table 4.3.15 shows that the predictions drawn from position in turn are confirmed: plosives – articulatorily stronger than fricatives – occur more frequently in initial position and the least on medial position; of all the positions in the word, the one that favours the use of the fricative the least is initial position. The results also tested to be statistically significant (*p-value = 5.923967 x 10^{-34}; Chi-square 153.02*). The hierarchy of use for position is middle > final > initial. Figure 4.3.2 illustrates these results.



These results confirm the results obtained in previous studies that suggested that in middle and final positions the fricative variants show higher rates of use, as opposed to the use of plosives in initial position. The examination of lexical item in interaction with position will help to reveal in greater detail the use of both variants. Table 4.3.15 shows this interaction:

Table 4.3.15 Cross-tabulation of lexical item and position, for fricatives only (N=1341)			
Position	Lexical Item	N	%
Initial	*Them*	79	81
	There (fore)	82	72
	Their	22	71
	Those	26	63
	That	94	61
	Than	24	60
	Others (collapsed)	32	59
	The	96	57
	This	42	55
	They	53	42
Middle	Other (collapsed)	124	98
	Other (-wise)	110	98
	Another	48	98
Final	*With*	121	81

From Table 4.3.15 we learn that whereas in middle position the use of fricatives is near categorical, initial position has a diverse behaviour going from 81% of use of the fricatives in *them*, to a 58% of preference of plosives in *they*. To some extent, this confirms that, first, intervocalic position has a strong effect on the use of fricatives (here near categorical) and, second, that preceding segment may be a stronger conditioning factor than following segment.

4.3.3.3.2.4 Summary of findings - linguistic constraints

The findings for linguistic constraints are:

- In preceding phonetic environment the hierarchy of use is vowels > consonants > pause. The preferred variant is the fricative, as predicted. Preceding pause showed similar rates of use for both variants.
- In following phonetic environment, a fricative is more likely to occur before a vowel than a consonant.
- The interaction of preceding and following segment confirms that intervocalic position strongly favours the use of fricatives.
- Fricatives are more frequent in turn initial position, whereas plosives are not particularly preferred in other positions.
- For position within the word it is middle position that favours the use of fricatives; the rates of use for plosives are, however, higher in initial position.

4.3.3.3.3 Multivariate analysis

The analysis of the combination of factors is required to uncover any possible patterns of interaction. I will consider all significant results in the factor-by-factor analysis. The table below displays the multivariate analysis for the majority group of variants: fricatives.

In a preliminary run of the MVA, I considered for each variant all factors examined previously, plus those that were not included in the analysis, to confirm that they are not statistically significant in the variation of voiced dental fricative. Nevertheless, to avoid the "kitchen-sink effect", I run in the MVA only those factor groups that seem to have a real effect on the variation of the examined feature; only internal factors are considered: position of the variable in the word and in turn, lexical item, and preceding segment; these results are displayed below.

Table 4.3.16 Multivariate analysis for voiced dental fricative (ð), for the majority group of variants only: fricatives

Log likelihood	-812.239	Input	.71
	FW	N	%
Position of the variable in the word			
Middle	.93	288	98
Final	.54	161	81
Initial	.30	903	61
Range	**63**		
Preceding phonetic environment			
Vowel	.72	585	88
Consonant	.38	495	63
Pause	.25	272	49
Range	**47**		
Lexical item			
Other, otherwise	[.94]	112	98
Another	[.94]	49	98
Others (<30)	[.66]	181	86
With	[.56]	161	81
Them	[.56]	98	81
There, therefore	[.44]	14	72
Their	[.43]	31	71
Those	[.35]	41	63
That	[.32]	155	61
Than	[.31]	40	60
The	[.29]	168	57
This	[.27]	76	55
They	[.18]	126	42
Position in turn			
other	[.51]	1293	72
initial	[.25]	59	46

Table 4.3.16 indicates the contribution of the factors to the application of the rule for fricatives. The input indicates that the use of fricatives is favoured (.71). Following the recoding of the factors found to be relevant in the factor-by-factor analysis (Young and Bayley, 1996), I included in this table only those factors found to be significant: position of the variable in the word and in turn, lexical item, and preceding segment. The range indicates the importance of the factor to the overall contribution of the application of the rule, that is, the higher the range, the greater the contribution (cf. previous multivariate analysis for more references). The examination of each of the significant groups, from the one that contributes the most to the least in the conditioning of voiced dental fricative (ð), suggests that:

- Position of the variable in the word was found to be the most significant factor, contributing the most to the use of fricatives (range 63); the positions that favour the application of fricatives are: middle (.93) and final (.54).
- The second factor group found to be significant is preceding segment, with ranges of .47. The hierarchy of use for the application of fricatives is vowels > consonants > pause, with only vowels favouring their use (.72).

The results, thus, suggest that only internal constraints are the source of variation in the use of voiced dental fricative by non-native speakers of English.

4.3.3.3.4 Summary and conclusions

The first finding obtained confirms the wide range of variation for voiced dental fricative (ð): the occurrence of nine different variants ranges from zero realization to five different forms of fricative and three forms of plosives. The five forms of fricatives are collapsed into a single group (fricatives) and the three forms of plosive also into one group (plosives). However, two variants are discarded from the analysis given the low number of occurrences (and their limited distribution): zero realization and the voiceless laminal – dentoalveolar fricative. The preferred group is "fricatives", as predicted.

The results drawn from the examination of the factor groups considered confirm that the variation shown by native speakers of Spanish in EFL contexts is similar to the variation exhibited by native speakers of English: the use of plosives is higher in initial position and of fricatives in middle and final positions, meaning that position in the word is one of the most important contributing factors in the variation of voiced dental fricative (ð). Against the predictions, external factors (sex, socioeconomic groups, and years of instruction in English) did not test significant. In relation to the linguistic factors, besides position in the word, the results for preceding segment suggest that vowels favour the use of fricatives and consonants and pause of plosives, as described in the literature; the effect of word class could not be tested as most lexical items examined were closed-class words.

The results obtained from the multivariate analysis confirm the effect of extralinguistic (individuality) and linguistic constraints: position of the variable in the

word and preceding segment were found to be significant in the application of the rule for both affricates and plosives. The discussion of the findings is completed in Chapter 5.

4.4 Postvocalic (r)

4.4.1 The variable

The study of this variable in particular has been considered one of the foundational variationist research topics since Labov's Martha's Vineyard study (1963). This variable shows different realisations not only between dialects (e.g., AAVE, RP and GA), but also within dialects, as was shown in Labov's study on the social stratification of New York City (1966).

The selection of it in my study is justified as in many foreign language settings traditionally the accent of instruction adopted has been RP,[144] but the effect of, e.g., the media may have brought to the classrooms features from other accents and dialects, such as GA or Australian English.[145] To this, the potential effect of the L1 must be added, that is, as shown in the second variable studied (voiceless alveolar affricate), there are some features that are being transferred from Chilean Spanish to English; from this, it can be expected that postvocalic (r), a feature that also is found in Spanish (with a different phonetic realisation from RP and GA), is very much likely to vary, given the wide range of possibilities offered, in terms of pronunciation, for EFL learners. Finally, this feature has been the subject of numerous variationist studies, which confirms its validity as an interesting feature to be looked at in this particular context (EFL).

4.4.2 Relevant findings and hypotheses

The variation of postvocalic (r) has been extensively studied in variationist research. One of the most important studies is undoubtedly Labov's 1966 New York Department Stores study, which set the basis for other variationist studies in terms of

[144] See Section 3.2.1.1.2 for more on the accent of instruction.

[145] For instance, Durham (2006: 61) suggests that in the case of the variation of English in Swiss speakers, "although American English may provide a strong linguistic model through the media", the preferred accent of instruction is Standard British English.

(1) identifying the variable as a "social differentiator in all levels of New York speech" and (2) demonstrating that the methodology utilised can be applied to any "systematic study of language" (Labov 1972: 44). Other important studies on postvocalic (r) are described below, both in native English and Spanish and in ESL/EFL contexts.

This section reviews the different research studies that have considered postvocalic (r) as a sociolinguistic variable. To determine if the L1 of the speakers in my sample is a factor that influences the realisation of the variable, I also explain the variation of postvocalic (r) in Chilean Spanish. In the analysis section, I will describe the use of the variable according to social and linguistic factors such as gender, years of instruction in English, socioeconomic group, lexical item, phonetic environment, and position of the variable within the word, as with the other variables. Finally, I will present the results in relation to the previous research available.

4.4.2.1 Variation in English

Postvocalic (r) presents in the different varieties of English and in different statuses. For instance, as reported by Romaine (1978: 144), this variable has been the subject of fluctuation due to prestige in English-speaking countries, varying from its loss from some British accents (considered an innovation or change from above) to its inclusion in some varieties in the U.S. Thus, whereas rhoticity is characteristic of General American (GA), non-rhoticity is in Received Pronunciation (RP).

As a sociolinguistic variable, postvocalic (r) has been studied in different settings and geographical areas. Labov (1972: 44) claims that this variable "appeared to be extraordinarily sensitive to any measure of social or stylistic stratification". In Martha's Vineyard, the greater degrees of retroflexion were linked to class and age, with local younger speakers showing higher degrees of retroflexion than non-locals. In the department stores study, Labov correlated the varying uses of postvocalic (r) to race, occupation, estimated age and style. Labov's results show that the variable (r) is a clear indicator of social stratification in terms of the extra-linguistic factors associated with the speakers, particularly race and occupation, and also to style, with emphatic style favouring the absence of postvocalic (r).

Labov (1972: 13-14) also argues that the variation of postvocalic (r) occurs in different dialectal zones of the U.S.; while it is generally regarded as rhotic, some areas – such as eastern New England, New York City and the South (Upper and Lower) – are non-rhotic. However, New York City presents its own variation, especially in terms of social factors, as pointed out earlier. Wolfram (1969) tested for the variation of postvocalic (r) in Detroit, where he correlated the presence or absence of the variable to sex differences over class, with women being more r-full than men. The effect of style has also been tested by Levine and Crockett (1966) in Piedmont; Shuy, Wolfram and Riley (1967) in Detroit; Feagin in Anniston (1990); Schönweitz in eight Southern states (2001);[146] Irwin and Nagy in Boston (2007); and Piercy (2012) in Dorset. While Wolfram (1969) suggests that – besides socioeconomic group, age, gender and style (factors considered by Labov) – racial isolation is a relevant conditioning factor,[147] Schönweitz (2001) has determined that rhoticity (in this case the use of retroflex /r/) is more commonly preferred by female, white, younger, highly educated, middle-class speakers.

In general, GA is considered a rhotic variety in the North and non-rhotic in the South. Some of the common realisations of /r/ found in the US are:

- Alveolar or post-alveolar approximant [ɹ]
- Retroflex approximant [ɻ]
- Weakly retroflex
- Vocalic deletion of /r/ (non-rhotic accents)
- A flap allophone of post-stress, pre-syllabic alveolar stops

(Ladefoged and Maddieson, 1996; Schönweitz, 2001; Ladefoged and Johnson, 2011)[148]

Despite the regional differences, rhotic GA is often referred to as a "network accent" (as it is frequently heard in media), and it is commonly linked, by NNSs, to speakers

[146] Schönweitz (2001) uses data from the *Linguistic Atlas of the Gulf States* (LAGS) to examine the variation of postvocalic /r/ in eight Southern states: Tennessee, Georgia, Florida, Alabama, Louisiana, Mississippi, Arkansas, and the eastern half of Texas.

[147] He distinguishes residential (individuals concentrated in certain neighbourhoods), educational (they attend certain schools, with black, white or mixed race predominance) and peer group isolation; in this case, isolation is understood as segregation.

[148] For other accounts on the different realisations of /r/ in American English see, e.g., Boyce and Espy-Wilson (1997); Guenther *et al.* (1999); and Kovecses (2000).

from the U.S. (with no particular regional differentiation); as such, it holds similar social appreciation to RP as a prestigious accent, and it is increasingly becoming popular in many EFL programmes. Some direct implications of this variety on my study are that, together with its increasing popularity, GA shares rhoticity with the L1 of the speakers, which means that its phonetic realisation is closer, or even more familiar, to the speakers' phonetic inventory than non-rhoticity. In practical terms, this familiarity may have an effect on speakers' fluency and, at the same time, it offers a wider range of target or native-like forms.

In the British Isles, the variation of postvocalic (r) is also present: according to Hughes, Trudgill and Watt (2005) postvocalic [ɹ] is preserved in central Lancashire, Ireland (Dublin), and Scotland. In Bristol, and generally in the southwest, postvocalic [ɹ] is quite retroflex, being produced as [ɻ], as in the speech of Belfast. In South Wales, "there is no post-vocalic [ɹ], except in the speech of some native speakers of Welsh; [ɹ] is normally an alveolar tap [ɾ]" (84); non-rhoticity is prestigious in the British Isles (see also, Trudgill, 2000).

As for the linguistic constraints involved in the variability of postvocalic (r) across accents, the effect of preceding vowel has been found to be highly significant in Boston and in Alabama, with different patterns, however, given different regional differences in vowel quality (Irwin and Nagy, 2007 and Feagin, 1990, respectively), and in Dorset (Piercy, 2012) just as word boundaries, with higher rate of r-less instances in non-final contexts (Labov, 1972; Irwin and Nagy, 2007; Piercy, 2012) and lexical versus function words, with the firsts favouring r-full instances over function words (Irwin and Nagy, 2007; Piercy, 2012). Piercy (2012) also considered stress (favouring the use of (r)), and lexical frequency (more frequent words disfavour the use of (r)).

As with many other phonetic variables, the realisation of postvocalic /r/ exhibits a continuum of uses. For example, Stuart-Smith (2007) notes four different groups of realisations of postvocalic /r/ in Glaswegian speakers,[149] Romaine (1978: 146)

[149] Stuart-Smith (2007: 1450), includes as articulated realizations of (r), the variants [ɻ], [R], [R], and [r], plus a group of pharyngealized and/or uvularized vowels, a group of 'plain' vowels, and a group of vowels followed by [h]/[ħ].

reports three groups of realisations for postvocalic /r/,[150] and Melchers and Shaw (2003: 51) discuss "several types of /r/ in England: [...) a weakly articulated post-alveolar approximant [ɹ], tapped [ɾ], especially in medial position, and increasingly a [w]-like type approximant. In Northumberland, uvular /r/ can still be heard". Thus, as shown here for diverse dialects of the British Isles, the varying continuum of realisations of postvocalic /r/ range not only in terms of manner of articulation, but also in place.

In terms of place of articulation, Cruttenden (2001: 28-30) shows /r/ is realised as post-alveolar ("as at the beginning of English *red*") and as retroflex ("such as is found in south-west British and some American English pronunciations of *red*"). In terms of manner of articulation, Cruttenden (2001) describes *r* as both tap ("as in many Scottish pronunciations of English /r/") and approximant ("the usual pronunciation in RP at the beginning of *red*").

Cruttenden (2001: 85) asserts that whereas in RP /r/ occurs only before vowels, in GA it can also occur before consonant and before pause; in Scottish it is replaced by the tap, although this also varies, alternating with [ɹ] being the latter commonly used (and more prestigiously) in postvocalic positions, and in London English, /r/ is replaced by [ɥ] or [ɯ̹].

The study of postvocalic (r) in other varieties of English includes Bayard (1995a) Starks and Bayard (2002) and Bauer and Warren (2004) in New Zealand English, Boberg (2004) in Canadian English, and Horvath (2004) in Australian English.

From the above-mentioned studies, we learn that the variation of postvocalic (r) is not only conditioned by external and internal factors, but that this occurs in different accents and dialects of English, which provides learners of EFL with a wide range of possibilities of target forms, depending on the accent of instruction, the prestige of the variable in the chosen model, or even their preferred TV shows.[151]

4.4.2.2 Variation in Spanish

[150] The first group is made up of [ɾ] and [ɽ], the second of [r] and the third of ∅, indicating zero realisation.

[151] As mentioned above, it is out of the scope of my study to assess whether preferences of the speakers are in fact a contributing factor to the variation of postvocalic (r), of what is the real effect of media on their EFL accent.

As this research is carried out in Chilean speakers of English, it is also important to consider the variation of postvocalic (r) in Chilean Spanish (recall that one of the general hypotheses formulated for my study suggests the transfer of phonemes from Chilean Spanish to English). In Spanish, this particular feature has not received the same attention as in English, and it has been mainly studied in relation to linguistic constraints.[152] As we see in Figure 4.4.1, in both final and medial positions it presents three allophonic realisations:[153] [ɾ], [r] and [˚ɹ̝] (Borland-Delorme, 2004).

/ɾ/	→[ɾ]~[r]~[ɹ̝] / _## _$/s/	"leer" "comercial"	[leer]~[leeɾ]~[leeɹ̝] [komeɾsjal]~[komersjal]~ [komeɹ̝sjal]
	→[ɹ̝] / como segundo elemento de la africada [t͡ɹ̝]	"otra"	[ot͡ɹ̝a]
	→[ɾ]~[r] / /bpd̪t̪gkf/_ _$	"otra" "creo" "porque"	[ot̪ra]~[ot̪ra] [kreo]~[kreo] [porḳe]~[porḳe]
	→[ɾ] / V_V	"parásito"	[parasit̪o]

Figure 4.4.1 Allophonic distribution of the cultured speech of Santiago – Borland-Delorme (2004)

Figure 4.4.2 also shows the allophonic distribution of traditional (peninsular) Spanish contrasted to the Chilean-cultured speech of the city of Santiago.

Traditional Peninsular Spanish		Cultured Chilean Spanish (Santiago)	
/ɾ/	[ɾ, r, ɹ]	/ɾ/	[ɾ, r, ɹ̝]
/r/	[r]	/r/	[r, ɹ̝]

Figure 4.4.2 Allophonic variation of traditional Spanish (left) and of cultured speech of Santiago (right) – Borland-Delorme (2004)

As shown in Figures 4.4.1 and 4.4.2, the variation exhibited in diverse varieties of English is also present in Chilean Spanish as opposed to traditional Peninsular Spanish (based on Borland-Delorme's findings); also, the absence of postvocalic (r) has been registered in Murcian Spanish (Gomez-Ortín, 2004; Hernandez-Campoy and

[152] I presume that the lack of studies in Chilean Spanish is due to the little attention that has been paid to variationist research there, especially in connection with linguistic conditioning, which, I believe, is understood as purely phonetic phenomena.

[153] Borland-Delorme (2004) uses the term allophone to refer to the different phonological variants used by the speakers.

Jimenez-Cano, 2004), in connection to social class (higher rates of (r) absence in lower classes), and in final position.

The gap in research prevents me from formulating any particular hypotheses for the linguistic behaviour of the speakers of the sample based on the variation described in their L1, but the contrast between the phonetic/phonemic inventory of English and Spanish – drawn from the relevant literature – plus the results obtained from the analysis, could help me identify those variants that belong to each language, in case there are substantial articulatory differences between them. I now turn to examine the previous research in both immigrant and instructional settings.

4.4.2.3 Variation in ESL/EFL contexts

One of the most prestigious accents amongst the Englishes of the world is RP, also the language of instruction in most ESL/EFL settings.[154] This accent became popular through the channel BBC News, known later as However, the increase in access to mass media communication from different parts of the globe also implies greater access to other varieties of English. This is particularly interesting in the use of this variable, as now learners of EFL/ESL are not only exposed to RP as their primary instruction accent, but also to all different sorts of varieties that, in some cases, are more easily available than RP, such as, e.g., GA, Australian English, etc. GA holds a similar status, as the language of U.S. broadcasting and ESL/EFL instructional settings.

As RP and GA are usually "referred to as the two *reference accents*" (Melchers and Shaw, 2003: 16; emphasis in the original), I considered these two varieties in my own research as indicatives for the variation in native speakers, and also because RP is characteristic of instructional settings[155] – used in academia as the primary language of instruction in this specific English teaching programme – and GA is the language of the media, to which speakers of my sample are exposed in recreational/ informal

[154] See previous note.

[155] It seems that, during the last two decades, the influence of RP has lessened given that it is often linked to the British elite; its gradual replacement towards a less socially-marked variety has led EFL instructors to choose Standard British English as the preferred dialect, as reported in countries such as, e.g. Germany (Sauer, 2002). Wells (1997) argues that RP remains only as a name and that it is now commonly recognised as "Standard Southern British English" (SSBE), which is technically an evolution of the term RP, as once coined by Jones (1917). Based on Wells' explanation, I will keep referring to this variety, as used in the instructional setting where I got my sample from, as RP.

settings. Neither of these varieties are used in their naturalistic scenarios; the first is an artificial educational setting created to teach English to Chilean Spanish speakers and the second is based on the import of films, television series and music into a non-English speaking country, which does not require the direct interaction with the speakers. Hence, I will test whether the variation of postvocalic (r) reflects one of the two accents to which the individuals of the sample have access to, or if their L1 is being transferred to their L2.

RP is described as non-rhotic, i.e., its speakers do not produce postvocalic (r) in words such as *car, cart*. In ESL/EFL instruction Roach (2004: 62), for instance, recommends to the foreign learner the realization of (r) as a post-alveolar approximant [ɹ], stating that the distribution of (r) in the BBC accent occurs only before vowels.

In relation to its treatment in ESL/EFL settings, postvocalic (r) has been traditionally linked to descriptive and prescriptive literature (see, e.g., Wells, 1992; Roach, 2004), and, more recently, influences NNSs' attitudes to the different accents of English (see, e.g., van Den Doel, 2010).

From the sociolinguistic perspective, the studies in NNSs of English have tended to reflect the NSs literature, particularly in connection to the conditioning factors. For instance, Hiang and Gupta (1992) studied the use of postvocalic (r) in Singaporean English, where the variant is considered prestigious; their results showed higher rates of use in female speakers, which is consistent with many studies that confirm the effect of gender on the use of prestigious and non-prestigious variants.

Dickerson and Dickerson (1977) and Beebe (1980) studied the effect of style[156] on the variation of postvocalic (r) in Japanese and Thai learners of English, respectively; their results show a strong correlation of style and following segment to the use of target forms. Also in relation to segment, Bonnici (2010) examined contracted forms (*they're*) in Maltese English, in which the r-full variant is much more common than the r-less variant; he found that both preceding and following segment contributed to the variation of postvocalic (r), with following vowels highly favouring the presence

[156] Major (1992: 190) suggests that the effect of style is based on how "closely speakers monitor their speech: the more they monitor the greater the accuracy", which explains the similar linguistic behaviour of NSs and NNs.

of (r). Also in relation to linguistic constraints, Chand (2010) studied the variation of (r) in Delhi English linked to morphological independence (e.g., runner) which tended to favour the occurrence of r-less instances.

Beebe's work (1980) is particularly interesting as she defines the different realisations of (r) in NNSs, classifying them into English native/non-native to Thai, IL variant, and variants transferred from Thai into English (see a full description of these in Beebe, 1980: 447); her work is relevant here as it shows that the variation of postvocalic (r) is, in EFL contexts, as systematic in native speakers, and that this occurs in a wide range of variants – some transferred from the L1.

In sum, the main findings for L2 situations tell us that postvocalic (r) is constrained by internal and external factors. Within the first, the most relevant are following segment (Dickerson and Dickerson, 1977; Beebe, 1980; Bonnici, 2010) and style (Dickerson and Dickerson, 1977; Beebe, 1980). Within the social factors, gender seems to be one of the most relevant factors (Hiang and Gupta, 1992).

4.4.2.4 Hypotheses

Based on the literature on postvocalic (r) that suggests that the use of this variable is constrained by social factors such as social class, age and sex and by linguistic factors, in particular phonetic environment, I propose to test that the variation exhibited by NNS of English reflects the sociolinguistic conditioning of postvocalic (r) exhibited by NS of either GA or RP, and mentioned in the section above; that is, the variation of Chilean speakers of EFL may fluctuate between rhoticity – characteristic of GA – and non-rhoticity, one of the most marked features of RP and Standard British English; the rhoticity should fluctuate between the target realisations of (r), the alveolar (canonical GA) and the retroflex approximants. Given the prestige of RP as accent of instruction over other accents, the preference of non-rhotic features is expected; however, if rhotic variants are preferred, the effect of either GA or the L1 of the speakers may be stronger on the use of postvocalic (r): if the presence of rhotic features decreases with the time of instruction, it might be because of the effect of the L1 (the presence of a variant that mirrors the L1 of the speakers might be expected, especially in the initial stages of their acquisitional process); if this sustains or increases over time, it is by effect of a rhotic accent. The main question, then, in

relation to the use of (r) in postvocalic position is which pattern do the speakers of the sample adopt?

4.4.2.5 Exclusions

Despite other studies having excluded pre-vocalic /r/ and /r/ following mid-central schwa vowel nuclei (e.g., Labov, 1966; Myhill, 1988: 205 "unless there was a word boundary between the /r/ and the vowel"), no tokens were excluded following this criterion,[157] but all lexical items whose following phonetic environment is some form of /r/ were excluded, e.g., *there're*, *other road*, etc., as in these cases it is not possible to identify if the variable belongs to the first or the second instance, e.g., *ther<u>e're</u>, other <u>r</u>oad.*

4.4.3 Variants and distribution

4.4.3.1 Variants

Six variants are evident in the current data:

- [r1] alveolar flap: this feature corresponds to the Spanish "hará"a/a'ra/; however, it is also registered in several accents and dialects of English, such as RP, GA, Australian and New Zealand English, especially as an allophone for /t/ and /d/ in intervocalic position: "but now in third year /jɪər/ we don't have grammar as a separate subject" (Bernardo, third-year male).
- [r2] alveolar approximant: this is considered the canonical GA phoneme158 here" /hɪəɹ/: "first /fɜɹst/, in high school you should, be everyday (…)" (Paola, first-year female).
- [r3] zero: corresponds to the RP standard for "here"e(/hɪə/): "It's hm, or extremely cold or [r3] extremely hot" (Caterina, third-year female).

[157] For instance, Myhill (1988) justifies the inclusion of prevocalic (r) as it has been shown that recurrent phenomena such as coda deletion reverse in the presence of vowels.

[158] For practical purposes, the alveolar approximant is defined as canonical for GA, based on traditional descriptive literature. This author acknowledges that (r) is a feature that undergoes a considerable range of variation in the different accents of English and that, specifically for the U.S., rhoticity goes from prestigious to stigmatised even at local level (e.g., New York City). See Section 4.4.2.1 for an account of the variation of postvocalic (r) in English as L1.

- [r4] = weakened alveolar approximant: identified as an almost zero realisation: “he came here, and he wanted to stay here one year, he was here/hɪəɹ/” (Carola, fifth-year female).
- [r5] = retroflex approximant: this feature is also characteristic of American English, along with [r2].159 “The teachers /ˈti·tʃəɻz/ are very old for being teaching” (Carmen, fifth-year female). Despite the fact that it is considered a native-like feature, I will consider the alveolar approximant as the canonical and this a second target for GA.
- [r6] = weakened alveolar flap: “I have to practise more /mɔɾ/, I'm in first year” (Francisco, first-year male).

The fact that there are six variants identified implies that the range of variation not only covers the standard forms for RP [r3], GA [r2] and [r5], but also other forms found in the different varieties of English studied. Also, the Spanish alveolar flap [r1] is present in the data.

4.4.3.2 Distribution

In this section, I describe the frequency of use of the different variants for postvocalic(r). The following discussion is subdivided into social and linguistic factors; within the first group, I consider individual speakers, sex, years of instruction in English and socioeconomic group. Within the second I include lexical item, preceding and following segment and position. This section is concluded with a multivariate analysis that examines the interaction of all the considered factors together.

4.4.3.2.1 Distributional data description for postvocalic (r)

The examination of the variable includes all selected items, with the exception of those mentioned in 4.4.3.5. Initially, I looked at all possible contexts in which the

159 Ladefoged and Johnson (2011) tell us that while alveolar approximant is more commonly found in American English, most speakers do not have a retroflex approximant. Some studies describe the retroflex as one more of the variants used by American speakers (see e.g., in variationist research, Schönweitz, 2001), whereas some phonetics studies suggest that “/r/ is realized almost universally as the retroflexed central approximant [ɻ]” (Stahlke, 2006: 57), other categorise it as “retroflex” or “bunched” (Boyce and Espy-Wilson, 1997: 3753). See the previous note on why I made the decision to consider the alveolar approximant the canonical and this – despite its recurrent description as (one of) the most prominent articulation (s) for (r) – a secondary target–like feature.

variable could occur, but then I decided to reduce the number of occurrences as their high frequency could distort the results.[160] This procedure is common in variationist research (see, e.g., Wolfram 1969, 1993: 214; Tagliamonte and Molfenter, 2007). Once I had identified all possible occurrences,[161] I extracted up to ten tokens per varying context.[162] Thus, I examine a total amount of 1862 tokens, for which the overall distribution is presented next.

4.4.3.2.1.1 Overall distribution

Table 4.4.1 introduced the overall distribution of the six variants found; the two majority variants correspond to the GA variants: alveolar approximant [r2] (44.6%, N=830) and the retroflex approximant [r5] (31.6%, N=31.6%).

Table 4.4.1 Overall distribution of postvocalic (r) (N= 1862)

[r1]		[r2]		[r3]		[r4]		[r5]		[r6]		Total
N	%	N	%	N	%	N	%	N	%	N	%	
26	1.4	830	44.6	210	11.3	120	6.4	588	31.6	88	4.8	1862

Table 4.4.1 reveals that the most frequent variant is [r2], the alveolar approximant, which corresponds to the GA canonical. Surprisingly (as a higher percentage of (r) deletion ([r3]) was expected given that it is the canonical feature for RP, the preferred accent of instruction), the second variant being preferred is [r5], the retroflex approximant. The [r3] variant, the zero realization of (r), corresponding to the RP standard, only appears on third place; the three remaining variants present low percentages of use, including [r1], the Spanish standard alveolar flap. In sum, the overall distribution of postvocalic (r) reveals that the most relevant findings are the high percentages of occurrence of the GA variants as opposed to the RP variant.

[160] With the preceding three variables I considered all instances without selecting a maximum number per lexical item, as most of them showed low frequencies (with the exception of items *English*, *Spanish* and *teacher* for the first variable), so its reduction made no sense.

[161] The instances selected were: are, first, for, here, learn (learned and learning), more (anymore) or, other (another and others), there, understand (understandable and understanding), work (working and works, plus fireworks), year, and your.

[162] This does not guarantee, however, that all speakers show a minimum use of all the instances, meaning that, not all speakers use all words, or not all the instances are analysable, as justified in the exceptions section above.

I now turn to how these data are distributed across individual speakers. Table 4.4.2 below shows in detail the distribution of postvocalic (r) in the six variants for all speakers.[163] It is arranged from the highest to the lowest use of the majority variant ([r2], the alveolar approximant). The preferred variant per speaker is shown in bold.

Table 4.4.2 Distribution of postvocalic (r), all variants, all speakers (N=1862)

	[r1]		[r2]		[r3]		[r4]		[r5]		[r6]	
Speaker	%	N	%	N	%	N	%	N	%	N	%	N
Miguel	0	0	**71**	83	1	1	3	4	20	23	5	6
Cesar	0	0	**63**	81	2	2	2	3	32	41	1	1
Sebastian	0	0	**60**	52	9	8	8	7	7	6	16	14
Carola	0	0	**58**	82	16	23	4	5	22	31	0	0
Marcela	2	1	**54**	28	6	3	2	1	23	12	14	7
Bernardo	1	1	**54**	60	1	1	2	2	38	42	5	6
Carmen	2	2	**53**	56	10	10	17	18	9	9	10	10
Caterina	0	0	**51**	54	21	22	19	20	5	5	4	4
Consuelo	3	2	**49**	39	8	6	11	9	30	24	0	0
Paola	0	0	**48**	42	5	4	10	9	32	28	5	4
Francisco	0	0	38	27	9	6	6	4	**45**	32	3	2
Andres	1	1	37	43	10	11	4	5	**47**	55	1	1
Melinda	0	0	36	47	5	6	5	6	**54**	70	0	0
Cecilia	8	8	**35**	34	8	8	2	2	30	29	16	15
David	1	1	32	32	10	10	2	2	**48**	48	8	8
Josefina	0	0	26	32	**64**	80	10	13	0	0	0	0
Juan	0	0	23	24	4	4	7	7	**66**	70	1	1
Jaime	10	10	14	14	5	5	3	3	**61**	63	9	9
Total	**1**	**26**	**45**	**830**	**11**	**210**	**6**	**120**	**32**	**588**	**5**	**88**

In contrast to the first three variables, there are no categorical speakers with over 95% of use of one of the variants. However, most speakers (11 out of 18, 61%) prefer the use of the alveolar approximant. There is a second group of speakers made up of the 33% of the sample who prefer the retroflex approximant [r5]. The only speaker who stands out of the general group is Josefina who presents the highest rate of use of the [r3] variant (64%, N=80). Josefina is a third-year, female speaker and is the only informant that has attended English lessons outside Universidad del Bío-bío; she is a middle-middle-class speaker. For her behaviour, she will be excluded from the analysis, as her use of the variant [r3] might obscure the results from the general

[163] With the previous variables, the collapsing of the minority variants was justified mainly as per their behaviour; if this would have been too dissimilar, the results would have been obscured. Here I present the whole range of variants to validate the omission of two of the six variants. If I had collapsed the variants in articulatory terms, e.g., [r2] and [r3], I would be grouping two variants that present rather heterogeneous behaviour.

group; also, her use of three of the six variants is zero categorical, which means that she uses mostly target and standard forms ([r3] 64% > [r2] 26% > [r4] 10%).

Another speaker that could be considered exceptional is Jaime: he presents a quite spread use of the six variants and one of the highest percentages of use of the [r5] form. Jaime is a first-year speaker, local from SEG D. Most of his instances of use occur in variants [r5] and [r2] (61%, N=63 and 14%, N=14, respectively). This speaker is also excluded from the analysis (for more on the removal of categorical contexts, see Tagliamonte, 2006).

I also exclude from now on the variants [r1] and [r6] given the small amount of tokens, which results in a better data distribution.

The distribution of the four remaining variables, for the 16 non-exceptional speakers is displayed in Table 4.4.3 below

Table 4.4.3 Distribution of postvocalic (r), four remaining variants, non-exceptional speakers (N=1534)

[r2]		[r3]		[r4]		[r5]	
%	N	%	N	%	N	%	N
51	782	8	125	7	104	34	523

The overall distribution shows that initial hierarchy of use ([r2] > [r5] > [r 3] > [r4]) is maintained. These general results tell us that the [r3] variant is not as commonly used as initially hypothesised, despite being considered the canonical form of one of the most prestigious varieties in EFL (RP), used as the primary accent of instruction. The most important finding is that the two GA variants [r2] and [r5], which for the purposes of this research are considered target variants, are used extensively, suggesting that the accent utilised in the instructional setting is not as influential in the acquisition of the L2 as expected.

A further, more detailed analysis of the data will determine if the use of a particular variant can be correlated to the different social and linguistic factors selected in this study.

4.4.3.3 Factor-by-factor analysis

4.4.3.3.1 Social constraints

To test the effect of each of the factor groups considered in the study of postvocalic (r), and as with the previous variables, in this section I analyse the use of the four groups of variants according to the social factors: gender, years of instruction in English, and socioeconomic group. The study of the previous variables tested the effect of social class and years of instruction in English as important contributing factors, especially in the use of target forms. Unlike the previous variables where target and non-target forms were identified, all variants can be classified as target, as they occur in different varieties of English. The question here is which of the target forms is the preferred by this group of NNSs.

The description of the variation by individual speakers is examined next.

4.4.3.3.1.1 Individual speakers

The examination of individuality for this group of speakers in the previous variants has revealed that, despite being a group of similar characteristics (e.g., in terms of age, they all belong to a form of middle class, they all attend university), their individual behaviour has been shown to contribute to the variation of the feature under study.

Table 4.4.4 shows the distribution of the four variants for postvocalic (r) in all non-exceptional speakers. The table is arranged from the highest to the lowest use of the majority variant [r2]. The preferred variant per speaker is shown in bold.

Table 4.4.4 Distribution of postvocalic (r), four variants, sixteen non-exceptional speakers (N=1534)

	[r2]		[r3]		[r4]		[r5]	
Speaker	%	N	%	N	%	N	%	N
Miguel	**75**	83	1	1	4	4	21	23
Sebastian	**71**	52	11	8	10	7	8	6
Marcela	**64**	28	7	3	2	1	27	12
Cesar	**64**	81	2	2	2	3	32	41
Carmen	**60**	56	11	10	19	18	10	9
Carola	**58**	82	16	23	4	5	22	31
Bernardo	**57**	59	1	1	2	2	40	42
Caterina	**54**	54	22	22	20	20	5	5
Paola	**50**	41	5	4	11	9	34	28
Consuelo	**50**	39	8	6	12	9	31	24
Cecilia	**47**	34	11	8	3	2	40	29
Francisco	39	27	9	6	6	4	**46**	32
Andres	38	43	10	11	4	5	**48**	55
Melinda	37	47	5	6	5	6	**54**	69
David	35	32	11	10	2	2	**52**	48
Juan	23	24	4	4	7	7	**66**	69
Total	**51**	**782**	**8**	**125**	**7**	**104**	**34**	**523**

As displayed in Table 4.4.4, after the removal of the exceptional speakers, the two groups that emerged from the previous section remain, with a larger group (69% of the speakers) showing higher rates of the variant [r2] (the alveolar approximant) and a smaller group (31% of the examined sample) showing higher use of the retroflex approximant [r5]. From the first, there is a greater number of female than male speakers, which is reversed in the second group. From the table, it is difficult to distinguish any other pattern, for instance, in relation to speakers' years of instruction in English or their social class; therefore, I will start by analysing factor by factor, in an attempt to unveil any possible pattern of variation. Thus, speakers' sex is the first factor group to be examined.

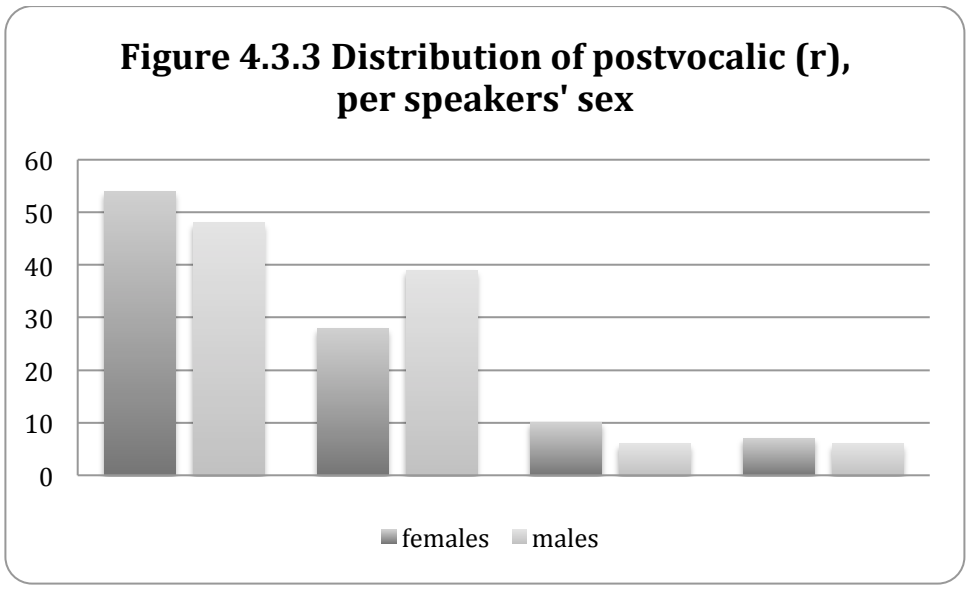
4.4.3.3.1.2 Speakers' sex

For the previous variables, sex has not been shown to be a particular important effect on the variation of Chilean Spanish speakers of English. In the case of postvocalic (r), the results suggest otherwise. Table 4.4.5 shows the distribution of the variable across the four variants for males and females.

Table 4.4.5 Distribution of postvocalic (r), four variants remaining, per speakers' sex (N=1534)

	[r2]		[r3]		[r4]		[r5]	
Sex	%	N	%	N	%	N	%	N
Females	54	381	10	72	7	52	28	198
Males	48	401	6	53	6	52	39	325
Total	**51**	**782**	**8**	**125**	**7**	**104**	**34**	**523**

From Table 4.4.5 we learn that both groups show the same hierarchy of use as in the OD; however, when examined from the variables perspective, there is an opposite pattern of use for the two majority variants, for males and females, with males using more [r5] than females. This pattern is better reflected in Figure 4.4.3.



The difference between males and females is statistically significant (*p-value=0.0.00002853023; Chi-square = 23.72*); when examined closely, however, both groups show considerable variation in relation to the intra-group behaviour, especially male speakers; therefore, how wise is it to group them together in function

of their gender? At least for females, whose range of percentile variation for the majority variant is 28, the pattern suggests that they behave similarly, unlike men, whose range for the use of the alveolar approximant variant is 52. This means that even the intra-group analysis shows that the behaviour of males is far more variable than females', that is, females do tend to behave similarly in their preference for the alveolar approximant variant; in other words, females approach the canonical GA variant more effectively than males, who vary heterogeneously across the variants, with an accent on the most rhotic element along the continuum.

4.4.3.3.1.3 Years of instruction in English

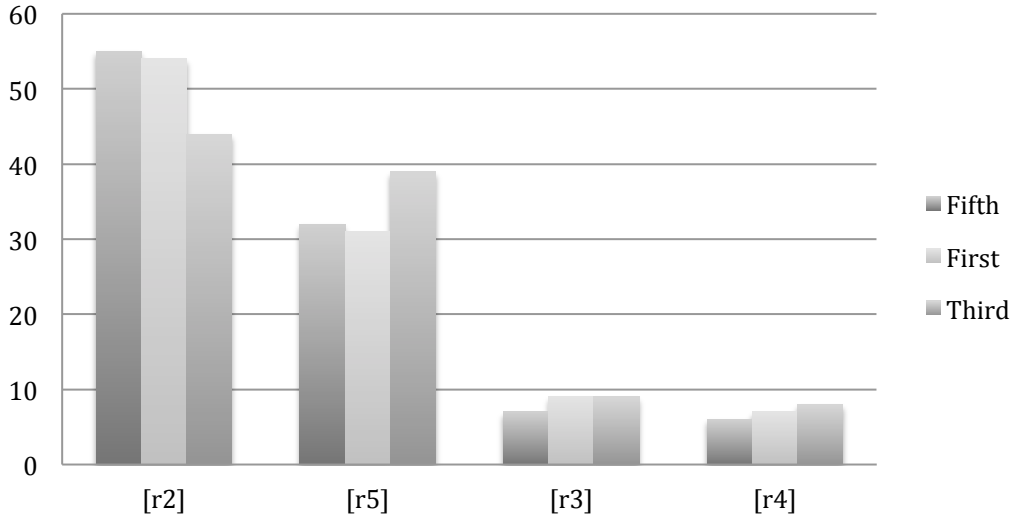
The examination of the previous variables determines that years of instruction in English, without being the most important factor accounting for the variation of Chilean speakers of English, needs to be considered as a source of variation, as, in general, more experienced speakers did show high rates of target forms. This, however, does not seem to be the case for postvocalic (r), as all three groups show similar rates for the four variants, as shown in Table 4.4.6 below.

Table 4.4.6 Distribution of postvocalic (r), four variants remaining, per speakers' years of instruction in English (YIE) (N=1534)

	[r2]		[r3]		[r4]		[r5]	
YIE	%	N	%	N	%	N	%	N
Fifth	55	392	7	53	6	41	32	228
First	54	182	9	29	7	23	31	107
Third	44	208	9	43	8	40	39	188
Total	**51**	**782**	**8**	**125**	**7**	**104**	**34**	**523**

The only great difference is the distribution of postvocalic (r) in third-year speakers, as the uses across the four variants seem a bit more spread, especially in relation to the two majority variants. These differences tested to be statistically significant (*p-value= 0.008766054; Chi-square = 17.15*) and are illustrated in Figure 4.4.4 below. Variables are arranged following the OD hierarchy, from highest to lowest.

Figure 4.4.4 Distribution of postvocalic (r) per speakers' years of instruction in English



Intra-group analysis revealed that, at least in relation to the use of the alveolar approximant [r2], the similarity in the behaviour of the speakers allows for them to be grouped together, but this is not the case for the second majority variant, especially in third-year speakers, which, to some extent, justifies the external variation, i.e., the highest range of uses of the variant [r5] are found in third-year speakers, who also show the highest percentage of use of the same variant. As all four variants are considered target forms in one or another accent, the acquisition of these can be considered as achieved, suggesting that instruction has no effect on the variation of this phoneme. I now turn to examine the variation from the socioeconomic group of the speakers.

4.4.3.3.1.4 Socioeconomic group (SEG)

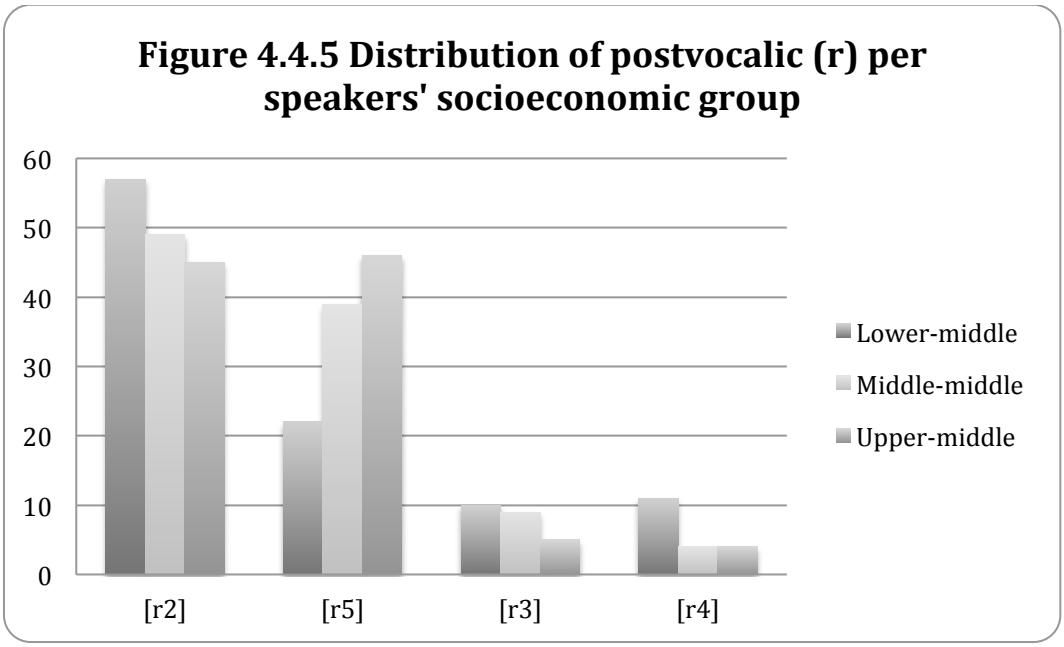
For this variable, socioeconomic class is considered a complex factor given that the prestige of the variants is relative, depending on the accent or dialect. It also has been stated that whereas [r3] is the most prestigious form in the accent of instruction of the speakers of the sample (RP), [r2] is the most widely spread form for postvocalic (r) through media (this corresponds to the GA canonical variant), along with the second majority variant [r5], which also has been identified as native-like (target) for speakers of GA. Table 4.4.7 shows the distribution of the four variants per socioeconomic group, in an attempt to show if the variants correlate to one of the

three social groups. Table 4.4.7 is arranged from the highest to the lowest use of the majority variant, shown in column 2 ([r2]).

Table 4.4.7 Distribution of postvocalic (r), four variants remaining, per speakers' socioeconomic group (SEG) (N=1534)

	[r2]		[r3]		[r4]		[r5]	
SEG	%	N	%	N	%	N	%	N
Lower-middle	57	311	10	52	11	61	22	121
Middle-middle	49	332	9	58	4	30	39	262
Upper-middle	45	139	5	15	4	13	46	140
Total	**51**	**782**	**8**	**125**	**7**	**104**	**34**	**523**

The results are particularly interesting for the lower- and upper-middle classes. Intra-group speaker analysis shows minor differences in the behaviour of the three groups, but this is quite stable, showing, for instance, that in the lower-middle class 5:6 speakers – and in the middle-middle class 5:7 speakers – prefer [r2], while in the upper-middle class 2:3 prefer [r5]. The lower class shows the highest rates of use of the alveolar approximant [r2] (57%, N=311), together with middle-middle-class speakers; the latter group, however, shows a closer difference with the retroflex approximant (49% and 39% respectively). In contrast, upper-middle-class speakers exhibit similar rates of use for both variants; this behaviour is displayed more clearly in Figure 4.4.5 below.



Again, the variability exhibited within the first majority variable is minimal, but this pattern is reversed in the second majority variant. But, what explanation can there be for this behaviour? Traditionally, in variables that are socially marked, higher rates of hypercorrection have been shown by the middle class than by upper-class speakers, but this is not the case, as, first, the most prestigious of the four variants is [r3] and this is barely used by the three class groups; and second, the fact must again be considered that all four variants are target forms, therefore the explanation is not in the achievement of target versus non-target variants. This factor group tested to be significant (*p-value= 4.624969 x 10^{-14}; Chi-square = 74.60*), therefore its behaviour will be examined again in interaction with other factor groups.

4.4.3.3.1.5 Interacting factors: years of instruction in English and socioeconomic group

As shown in the literature, the interaction of factors, and not the effect of one particular constraint, is frequently the source of variation (see, e.g., Labov, 1972, 2001). Two of the most relevant factors in the variation of postvocalic (r) are (1) years of instruction in English and (2) socioeconomic group; the interaction of these two factor groups is examined next.

Table 4.4.8 Cross-tabulation of years of instruction in English and socioeconomic class

			Years of instruction in English							
			First		Third		Fifth		∑	
			N	%	N	%	N	%	N	%
Socioeconomic class	Upper-middle	[r2]	0	--	56	28	83	75	139	45
		[r3]	0	--	14	7	1	1	15	5
		[r4]	0	--	9	5	4	4	13	4
		[r5]	0		117	60	23	21	140	46
		∑:	0		196		111		307	
	Middle-middle	[r2]	62	53	98	54	172	45	332	49
		[r3]	11	9	7	4	40	10	58	9
		[r4]	3	3	11	6	16	4	30	4
		[r5]	41	35	66	36	155	41	262	38
		∑:	117		182		383		682	
	Lower-middle	[r2]	120	54	54	53	137	62	311	57
		[r3]	18	8	22	22	12	5	52	10
		[r4]	20	9	20	20	21	10	61	11
		[r5]	66	29	5	5	50	23	121	22
		∑:	224		101		220		545	
	∑:	[r2]	182	54	208	4	392	55	782	51
		[r3]	29	8	43	9	53	7	125	8
		[r4]	23	7	40	8	41	6	104	7
		[r5]	107	31	188	39	228	32	523	34
		∑:	341		479		714		1534	

Starting with upper-middle-class speakers, the cross-tabulation of class and years of instruction in English suggests that the use of the alveolar approximant [r2] increases from third to fifth year (there are no first year speakers from upper-middle class). In the case of middle-middle-class speakers, the pattern for [r2] is more spread, with similar rates of use in first-, third- and fifth-year speakers. The percentages of use are also similar for the variant [r5]. Finally, lower-middle-class speakers also show similar rates of use for the alveolar approximant, but not for the retroflex approximant.

The results displayed in Table 4.4.8 suggest that there seems to be no direct relationship between the years of instruction in English and social class, probably with the exception of upper-class speakers, who show an increasing use of the [r2] variant. A similar increasing rate pattern is observed in middle-middle-class speakers, but for the retroflex approximant.

Having examined each of the selected factors individually, and then the interaction of two of the most relevant external factors, I now turn to the analysis of internal factors, but before that, I present a summary of the main findings for each social factor.

4.4.3.3.1.6 Summary of findings – social factors

The most important finding in relation to the use of postvocalic (r) is the occurrence of six different variants; from these, the four most used are examined: the alveolar approximant [r2], the zero realisation [r3], a weakened form of alveolar approximant [r4] and a retroflex approximant [r5]. From these, [r2] and [r5] account for almost 85% of the tokens.

The factor-by-factor analysis of social constraints, after the removal of two exceptional speakers (Josefina and Jaime – no categorical speakers found), revealed that:

Individual speakers

The variability of postvocalic (r) exhibited by individual speakers revealed two groups, the majority of whom (69%) preferred the use of the alveolar approximant over the alveolar retroflex. A preliminary examination of these individuals revealed that sex might be a relevant conditioning factor.

Sex

Gender differences suggest that both groups present the same hierarchy of use of the four variants; however, the use of the retroflex is higher in males than the alveolar approximants is in females.

Years of instruction in English

The increasing pattern of use of target forms from first to third year seen in previous variables is not shown here, as first, all variants are target forms. The majority variant [r2], however, which corresponds to the GA canonical, shows higher rates in all three groups, if compared to the second majority variant, also GA target.

Socioeconomic group

The data showed differences in the variation attributable to social class, as the three groups patterned differently, especially in relation to middle-middle-class speakers, who showed a more spread distribution and upper-middle-class speakers, who showed higher rates of the retroflex approximant. Nevertheless, the interaction of this factor group with years of instruction in English showed the weak effect of this factor to the overall variation of postvocalic (r). Hence the examination of internal factors is required.

4.4.3.3.2 Linguistic constraints

As with the preceding variables, the examination of linguistic constraints is carried out, as postvocalic (r) has shown to be particularly constrained by preceding and following segments, and position (Cucchiarini and van den Heuvel, 1999; also Nagy and Irwin, 2010).

4.4.3.3.2.1 Lexical item

In order to assess if the variation is determined by linguistic internal factors, the following lexical items were analysed: *other* (another), *there, for, more, are, your* (yours, yourself), *year* (years, year's), *or, first, here, work* (works, working), *understand* (understandable), *learn* (learns, learned, learning). Also a few lexical items containing the suffix –er were selected (as in *worker, teacher*); I also included the category *others* with tokens ending on –er/-or (as in *winter / professor*) with less than ten occurrences. Table 4.4.9 displays all individual lexical items, except the category suffixes (these are examined later).

Table 4.4.9 Distribution of postvocalic (r), four variants remaining, per lexical item (N=1393)

	[r2]		[r3]		[r4]		[r5]	
Lexical item	%	N	%	N	%	N	%	N
Your	77	47	8	5	8	5	6	4
First	74	54	8	6	1	1	16	12
Here	64	66	0	0	6	6	30	31
Work	60	44	6	4	7	5	27	20
Others	54	76	5	7	4	6	37	53
Year	54	55	6	6	9	9	31	32
For	49	60	19	23	8	10	24	29
More	48	50	12	12	1	1	39	41
Understand	48	21	16	7	23	10	14	6
Are	46	65	2	3	5	7	46	65
an(other)	44	45	15	15	8	8	34	35
There	43	47	8	9	11	12	38	41
Learn(-ing)	41	44	5	5	3	3	51	55
Or	35	38	19	21	16	17	31	34
Total	**52**	**712**	**8**	**123**	**8**	**100**	**30**	**458**

The preferred variant in almost every lexical item is [r2], with two exceptions: *are*, which shows similar rates as [r5], and *learn(-ing*), which shows almost 10% more of occurrences with [r5] than with [r2]. As individual lexical items, it is difficult to test for a general pattern. All the examined words are very frequent in English, especially in the case of function words, which account for about 50% of the uses of the preferred variant.

Now, if we consider the words above in grammatical terms, all of them can be considered (mono-) morphemes. As such, they can be contrasted with the suffixes found in the data. The following table displays the results for such a comparison:

Table 4.4.10 Distribution of postvocalic (r), four variants remaining, per lexical item (N=1534)

	[r2]		[r3]		[r4]		[r5]	
Lexical item	%	N	%	N	%	N	%	N
(Mono-) morphemes	53	712	9	123	8	100	30	456
Suffixes	50	70	1	2	3	4	46	67
Total	**52**	**782**	**6**	**125**	**6**	**104**	**38**	**523**

First, the chi-square test revealed that the differences in this factor group are significant (*p-value=0.00001569388; Chi-square = 20.16*). Second, the difference between the two most frequent variants ([r2] and [r5]) seems closer to suffixes than to (mono-) morphemes, which appears more spread across the four variants, suggesting that that the variation might be linked to the lexical items and not, for instance, to the position of the words (final). But, the internal behaviour of the elements of the group is dissimilar, therefore it is not a truthful picture of the variation of these lexical items.

Now, could the variation be justified in terms of word class? Table 4.4.11 below displays the distribution of content versus function words for the four variants.

Table 4.4.11 Distribution of postvocalic (r), four variants remaining, per word class (N=1534)

	[r2]		[r3]		[r4]		[r5]	
Word class	%	N	%	N	%	N	%	N
Content words	52	364	4	37	5	38	39	243
Function words	49	418	10	88	8	66	33	280
Total	**51**	**784**	**8**	**125**	**7**	**104**	**34**	**525**

As shown, content and function words seem to vary similarly, especially in the majority variants, which might suggest that the variation is not found in this group; however, the differences between the groups are statistically significant (*p-value=0.0011073; Chi-square = 16.05*). As I did for the previous variables, in the

following sections I examine the effect of segment (both preceding and following), position and word class.

4.4.3.3.2.2 Phonetic environment

4.4.3.3.2.2.1 Preceding phonetic environment

Following the methodology used by Nagy and Irwin (2010) in relation to the treatment of preceding vowels, I collapsed all vowels into six categories: [a], [e], [i], [o], [u], and schwa. In Nagy and Irwin's study (2010), the vowel that presented the highest rates of deletion of postvocalic (r) is schwa,[164] and the least was "o"[165] – which is also expected here; the distribution of the four variants for postvocalic (r) is displayed in Table 4.4.12 below.

Table 4.4.12 Distribution of postvocalic (r), four variants remaining, preceding segment (vowel groups) (N=1534)

	[r2]		[r3]		[r4]		[r5]	
Vowel groups	%	N	%	N	%	N	%	N
[u]	100	3	0	0	0	0	0	0
[i]	64	66	0	0	6	6	30	31
[e]	51	383	6	47	6	43	37	272
[o]	50	265	12	65	8	40	30	156
[a]	47	65	2	3	5	7	46	64
schwa	0	0	56	10	44	8	0	0
Total	**51**	**782**	**8**	**125**	**7**	**104**	**34**	**523**

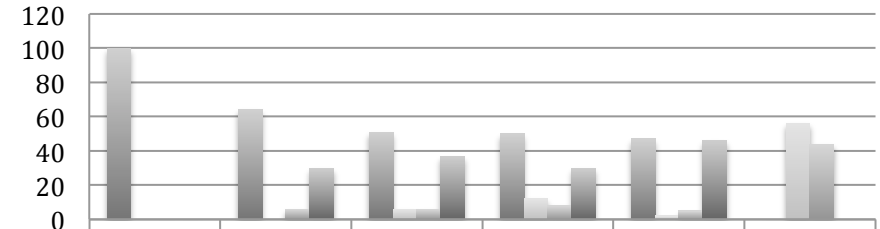
Of the six vowel groups, there seems to be a decreasing use of the majority variant (alveolar approximant) for the first five groups, depending on the openness of their articulation. The group [a] shows the same percentages of use for variants [r2] and [r5], and the clear exception to preceding segment is schwa, which presents the highest rates on the non-majority variants (zero realisation and weakened alveolar

[164] Nagy and Irwin (2010) make a distinction between stressed and unstressed schwa, the first favouring the deletion and the second disfavouring it. This distinction is not made in my study.

[165] The hierarchy shown by Nagy and Irwin (2010) for the deletion of postvocalic (r) is stressed schwa > a > e > u > I > o > unstressed schwa (cf. previous note).

approximant). The distribution of postvocalic (r) is better displayed across the six vocalic groups for the four variants studies in Figure 4.4.6 below.

Figure 4.4.6 Distribution of postvocalic (r) in preceding phonetic environment



	[u]	[i]	[e]	[o]	[a]	schwa
[r2]	100	64	51	50	47	0
[r3]	0	0	6	12	2	56
[r4]	0	6	6	8	5	44
[r5]	0	30	37	30	46	0

From Figure 4.4.6 above we learn that the openness of the vocalic groups might, effectively, contribute to the decreasing use of the alveolar approximant, whereas the rates of use of [r5] seem to be concentrated in middle articulatory positions. Finally, schwa shows a completely different pattern, which does not seem to be related to the openness of the vowel groups. The differences in this group were found to be significant (*p-value= 1.16655 x 10^{-24}; Chi-square = 151.59*); it is expected that similar results are found in following segment, as they would confirm the effect of articulation on the realisation of postvocalic (r).

4.4.3.3.2.2.2 Following phonetic environment

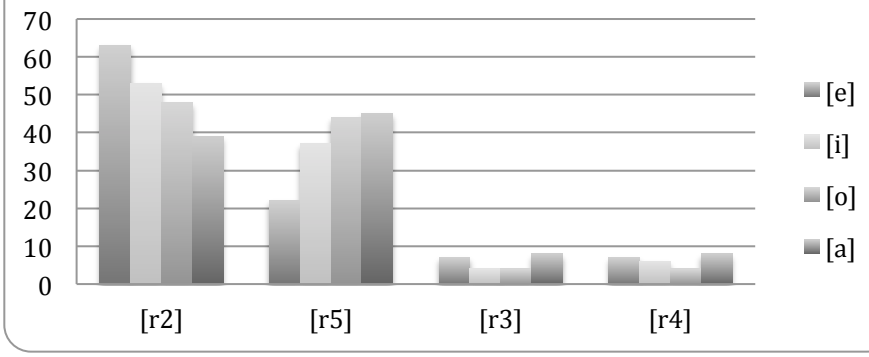
Nagy and Irwin (2010) propose that "the following segment constraint is universal in that following vowels (the linking ***r*** context) favour (r-1) more than following consonants or pause". Thus, a similar pattern is expected for following segment in the variation of postvocalic (r); I consider in the analysis two larger categories: vowels (four vocalic groups) and consonants (six groups and pause), as justified before for the other variables (see also, Labov, 1966; Cedergren, 1973; Nagy and Irwin, 2010). Table 4.4.13 displays the distribution of the four variants for the following segmental groups considered. In bold are shown the highest scores for the majority variants and in italics the scores that are similarly distributed in [r2] and [r5]:

Table 4.4.13 Distribution of postvocalic (r), four variants remaining, following segment (vowel groups and consonants) (N=1534)

	[r2]		[r3]		[r4]		[r5]	
Vowels	%	N	%	N	%	N	%	N
[e]	**63**	26	7	3	7	3	22	9
[i]	**53**	57	4	4	6	6	37	40
[o]	*48*	12	4	1	4	1	*44*	11
[a]	*39*	46	8	10	8	9	*45*	54
Consonants								
Fricatives	**57**	267	8	35	8	36	28	131
Plosives	**57**	144	9	24	7	18	27	68
Affricate	**50**	9	0	0	11	2	39	7
Lateral	*48*	11	4	1	0	0	*48*	11
Pause	*47*	100	7	15	5	11	*41*	86
Nasals	*42*	83	11	22	6	11	*42*	83
Approximants	*40*	27	15	10	10	7	*34*	23
Total	**51**	**782**	**8**	**125**	**7**	**104**	**34**	**523**

The differences in this factor group are statistically significant (*p-value = 0.002088987; Chi-square = 57.01*). As displayed in Table 4.4.12, the openness of following vowels seems to replicate the patterns found in preceding vowels, with alveolar approximant being found more frequently with closed vowels than back vowels. This decreasing pattern is displayed more clearly in Figure 4.4.7 below.

Figure 4.4.7 Distribution of postvocalic (r) in following phonetic environment, vowels only



In relation to consonant groups, the decreasing use of the alveolar approximant follows the pattern fricatives > plosives > affricates > laterals > nasals > approximants. This is consistent with the diminishing in sonority scale presented by Major (1996) (see Note 33, Section 2.1.3.2.2.1), which goes from +sonorant vowels to –sonorant stops. In this case, the use of the alveolar approximants seems to decrease in accordance with the sonority of the consonantic groups. The examination of preceding and following segment will be reviewed in Section 4.4.3.3.3, *Multivariate analysis* for postvocalic (r).

The final factor group examined in connection to linguistic constraints, more specifically with lexical item, is the position of the vowel in the word.

4.4.3.3.2.3 Position within the word

The literature suggests that the position of the variable in the word is an important contributing factor for the variation of diverse phonetic variables (see, for instance, the previous variables studied here). For postvocalic (r), it has been found that final position, when followed by a vowel or by a pause, favours highly the deletion of (r) (Nagy and Irwin, 2010). This phenomenon does not only occur in relation to postvocalic (r), but it can be considered a rather frequent phenomenon in some languages, such as Spanish (Alarcos-Llorach, 1958; Goldsmith, 1981). The distribution of postvocalic (r) in relation to the position of the variable in the word is shown in Table 4.4.14

4.4.14 Distribution of postvocalic (r), four variants remaining, per position of the variable in the word (N=1534)

	[r2]		[r3]		[r4]		[r5]	
Position	%	N	%	N	%	N	%	N
Middle	56	231	6	23	7	28	32	133
Final	49	551	9	102	7	76	35	390
Total	**51**	**782**	**8**	**125**	**7**	**104**	**34**	**523**

Chi-square test reveals that this group is significant (*p-value=0.04830559; Chi-square = 7.89*). Also, the results displayed in Table 4.4.14 suggest that this factor group does not seem to account for the variation exhibited: both middle and final positions present similar behaviour for the three variants.

Having examined all factors individually, I conclude this section with a summary of the findings for the linguistic constraints. The analysis of the variation of postvocalic (r) concludes with a run of multivariate analysis.

4.4.3.3.2.4 Summary of findings - linguistic constraints

After completing the examination of every individual factor, the main findings for the linguistic constraints suggest that:

- Most lexical items present high rates of use of the [r2] variant (alveolar approximant), with two exceptions: are, and learn(-ing).
- The analysis of word class reveals that, despite being statistically significant, the behavior of both function and content is very similar, therefore the effect of class does not contribute to the variation of postvocalic (r).
- For preceding segment, schwa is the preceding vowel that favours most the use of the RP variant ([r3]). There seems to be a correspondence between the articulatory properties of vowels in the use of either the alveolar or the retroflex variants (open and middle articulation, respectively).
- For following segment two patterns emerge: first, the selection of the alveolar or the retroflex seems to replicate the results for preceding segment in relation to the openness of the vowels; second, for consonants, the articulation of the features follows seems to be linked to the sonority of the following segment (+sonority for the alveolar, - sonority for the retroflex).

- The position of the variable in the word is barely significant and the results are very similar. This factor does not seem to contribute to the varying use of postvocalic (r).

To test whether the interaction of factors is the source of variation of postvocalic (r), I run a multiple regression analysis on all relevant factors. The results are presented in the next section.

4.4.3.3.3 Multivariate analysis

As with the previous variable, in order to conduct multivariate analysis, I collapsed the variants into groups to avoid the occurrence of empty cells. I also excluded from the analysis two categorical contexts in the group factor *lexical item*, as they do not exhibit variation. The results for the application of the variable rule for the two majority variants are shown in Table 4.4.15 below; the results are arranged from highest to lowest factor weight. Only the groups found significant in the factor-by-factor analysis are included; the groups years of instruction in English, individual speakers, individual lexical items, word class, preceding segment and position are not examined.

Table 4.3.15 Multivariate analysis for postvocalic (r), for the two majority group of variants: alveolar approximant [r2] and retroflex approximant [r5]

	[r2]		[r5]		
	Input	.51	Input	.34	
	Log likelihood	-1062.994	Log likelihood	-984.301	
	%	FW	%	FW	N total
Socioeconomic group					
Lower-middle	57	.56	22	.36	545
Middle-middle	49	.48	38	.55	682
Upper-middle	45	.44	46	.63	307
Range		**12**		**27**	
Sex					
Females	54	.53	28	.43	703
Males	48	.47	39	.56	831
Range		**06**		**13**	
Morphological category					
(Mono-) morpheme	51	[.50]	33	.49	1391
Suffix	49	[.48]	47	.63	143
Range		--		**14**	
Following segment					
Vowels					
[e]	63	[.63]	22	.36	41
[i]	53	[.52]	37	.54	107
[o]	48	[.47]	44	.61	25
[a]	39	[.38]	45	.62	119
Consonants					
Fricatives	57	[.56]	28	.43	469
Plosives	57	[.56]	27	.42	254
Affricates	50	[.49]	39	.56	18
Laterals	48	[.47]	48	.64	23
Pause	47	[.46]	41	.57	212
Nasals	42	[.41]	42	.58	199
Approximants	40	[.39]	34	.51	67

The results from the multivariate analysis suggest:

- The most significant group for both variants is *SEG* (range 12 for alveolar approximant and 17 for retroflex approximant), which suggests that class is the most relevant conditioning factor.
- Whereas the second contributing factor for the majority variant [r2] is sex (range 06), for [r5] is third (range 13). Intra group behaviour, however, determined that whereas men spread across the variants, women showed a more consistent behavior on the use of the alveolar approximant.

- The only linguistic factor found significant in the variation of postvocalic (r) is morphological category (range 14), for the variant [r5].

The effect of the selected factors is discussed with greater detail in the following chapter.

4.4.3.3.4 Summary and conclusions

The most important finding in relation to the use of postvocalic (r) is the occurrence of six different variants: ranging from zero-realisation to a full flap; the analysis of these is reduced to the four most used variants: [r2] an alveolar approximant, [r3] zero realisation, [r4] a weakened alveolar approximant, and [r5] a retroflex approximant. All the variants are considered target, as they are present in many dialects and accents of English. For practical purposes, [r2] is treated as GA canonical, [r3] as RP, and [r5] as GA target.

The preferred variant is the alveolar approximant [r2] followed by the retroflex approximant [r5]. The other two variants account for only 15% of the total instances.

The general hypothesis proposed – the similarity between the sociolinguistic behaviour of native and non-native speakers – was confirmed, suggesting that non-native speaker variation behaves similarly to native variation: postvocalic (r) is constrained by internal and external factors, as described below for the secondary hypotheses:

For sex, intra-group variability showed that only the behaviour of females was consistent in the use of the alveolar approximant variant, which is the majority variable for most speakers. This means that only a few males are responsible for the high rates of use of the second majority variable, retroflex approximant.

For years of instruction in English, no relevant patterns were uncovered. As stated before, all four variants are considered target forms, therefore, in acquisitional terms, all individuals are performing native-like, with a clear preference for the GA standard form.

Socioeconomic group tested to be significant, in both inter- and intra-group behaviour. With upper- and lower-middle-class speakers showing dissimilar

linguistic preference for the two majority variants, middle-middle-class individuals were the most variable among the four variants.

The final hypothesis concerned the constraining of preceding and following segments and the prosodic position of the variable in the word, which shed some light in terms of the effect of openness of the vocalic context and the sonority of consonants – phenomena which have both been found to be consistent with the literature on NSs of English variability.

The discussion of all results is presented in the final chapter, which follows.

CHAPTER 5 – DISCUSSION AND CONCLUSION

5.1 Introduction

Despite the vast number of studies on variation that followed the seminal work by Labov (1963, 1966), especially in L1 situations, and more specifically in English dialects and accents, it is only recently that variationist research has turned to the study of second-language. Nonetheless, the study of non-immigrant situations (e.g., Regan, 1996; Major, 1996; Durham, 2006) is comparatively scarce for immigration settings (see e.g., Poplack, 1978 and Wolfram, 1971; Sankoff *et al*., 1997; Mougeon and Rehner, 2001; Rehner, Mougeon and Nadasdi, 2003; Nagy, Blondeau and Auger, 2003).[166] With my work, I look to contribute to the study of variation in foreign language situations. But, is it possible to explain the variation of this group of speakers based on their linguistic performance on the four variables examined? Furthermore, is it possible to determine if said variation is the result of the interaction of the two languages (English and Spanish) or the result of the interaction of social and linguistic factors? Is this variation similar to that exhibited by native speakers of English, or is it characteristic of Chilean speakers of EFL?

Recall that the aims of this study are:

I. to establish whether there are systematic patterns of variation in NNS in the context of EFL, just as there are in native speech and non-native immersion contexts;
II. to identify potential similarities and differences between the patterns of variation exhibited by NS and NNS; that is, to test whether the variation in NS of English is similar or different in NNS, and to contrast the factors involved in the variation for these two groups of speakers.

As part of the second aim, I looked to test whether the patterns of variation for Chilean speakers of EFL were transferred to English. The prior relevant literature has only been able to partially answer these research questions. For instance, in relation to the transfer of sociolinguistic patterns, Beebe (2006: 433) suggests that

[166] Most of these studies, with the exception of the first two, focus on morphosyntactic and discourse variation in French-speaking settings as L2.

"interlanguage becomes permeable to a superordinate rule system in formal situations." More specifically, the second-language system, in this case English, is able to act as "the superordinate rule system when the variable examined had no social meaning in the native language"; if the variable is marked for social value in the L1, this process reverses, i.e., the L1 is "adopted as the superordinate rule system". This was to some extent tested for one of the variables: with voiceless alveolar affricate, despite the fact that the majority variant is the canonical /tʃ/, the occurrence of the prestigious Chilean form is documented in EFL. This study shows that the four variables selected have different conditioning factors, but what is the common ground behind the variation of this group of individuals? In the next sections, I examine each of the hypotheses proposed with the main results per variable.

5.2 Testing the hypotheses

5.2.1 Hypercorrection occurs as a result of the transfer of prestigious variants from Chilean Spanish to English.

This hypothesis is refuted for most variables. For all four features, the majority variant was the target or canonical form, which means that there is no transfer of Spanish forms into English. A possible exception is, as mentioned above, voiceless alveolar affricate, which showed a significant percentage of use of the Chilean Spanish prestigious form that is not present in the English phonetic inventory. Another feature whose behaviour might be interesting to observe in terms of transfer is voiced dental fricative, as it shows similar conditioning in English and Spanish;[167] the difference is, however, in relation to the conditioning factors, and whereas for voiceless alveolar fricative the effect of SEG and sex are relevant, the conditioning of voiced dental fricative seems to be motivated by linguistic factors, therefore it cannot be considered a case of hypercorrection, at least in Labov's terms.

5.2.2 Males and females pattern differently

Females and males did pattern differently in two of the variables: voiceless alveolar fricative and postvocalic (r). In the first, however, the differences between males and females are not significant; females tended to use the target forms over males, and in

[167] Recall that in Spanish this is an allophone for /d/, occurring mainly in intervocalic position.

the second, only the females exhibited a consistent intra-group behaviour, favouring more markedly the use of the target alveolar approximant over the (also target) retroflex approximant. I expected a higher effect of sex on the variation of the four features, given the role of gender in variationist research (see, e.g., Labov, 2001), but my study tests that sex differences are not as relevant for this particular group of individuals in this specific context (non-immigration, instructional), as it was for Hiang and Gupta (1992). I will come back to sex differentiation later.

5.2.3 The phonetic environment is a relevant contributing factor in variation

As for the phonetic environment, the hypothesis is also confirmed, although partially. Preceding segment showed consistencies with the literature for voiceless alveolar fricative, voiced dental fricative and postvocalic (r), in L1 and L2 settings. Following segment also showed consistencies with the literature for voiceless alveolar affricate, voiced dental fricative and postvocalic (r). The results from the tests for significance suggest that phonetic environment has an effect on the variability of the aforementioned variables (this is in the factor-by-factor analysis); some of these general results suggest that intervocalic position tends to contribute reliably to the preference of weaker forms in articulatory terms (that is, in the sonority hierarchy suggested by Major, 1996). The restraining of stronger articulations in consonantal contexts (as suggested by Cedergren, 1973) was not particularly tested, however.

For the voiceless alveolar fricative, the tests revealed that preceding and following segments contribute statistically to the variation; the results suggest that, as tested by Cedergren (1973), the occurrence of preceding vowels increases the possibility of choosing a fricative over an affricate, as found in my study. In articulatory terms, this is explained by lenition, a common phenomenon in which the articulation of consonants is weakened mainly by the effect of the sonority of adjacent vowels. The same articulatory explanation can be provided for the effect of following segment in voiceless alveolar affricate (vowels preferring affricates over stops, the latter articulatorily stronger than the firsts), of voiced dental fricatives (with highest rates of fricatives over plosives in intervocalic positions) and of postvocalic (r). In this last variable, further than the general effect of vowels on the articulatory lenition of segments (as all preceding context are vowels), the analysis reveals that the selection of either the alveolar of the retroflex variants will depend on the openness of vowels,

with front vowels favouring the use of the alveolar and back vowels of retroflex variants. Thus, the effect of segment is not only consistent with variationist research, but also with articulatory phonetic phenomena (see, e.g., Labov, 1994; Ladefoged and Maddieson, 1996; Ladefoged and Johnson, 2011).

5.2.4 The amount of exposure to English in formal education is equivalent to the use of target forms

Perhaps the most robust constraining factor on the variation of most variables is the amount of time the speakers of the sample have been exposed to English by means of formal instruction: the years of instruction in English are consistently proportional to the use of target forms. This is not only relevant in variationist terms, but also for the acquisitional implications of approximation to native-like performance (I return to this point in the implications section).

From the analysis we learn that the amount of time the speakers have been exposed to English in the instructional setting is statistically significant for the variation of voiceless alveolar affricate, voiced dental fricative and postvocalic (r). In relation to the first variable (non-significant), the increasing use of the target form (fricative) with the proportional abandonment of the Spanish (standard) native-like forms reveals that the pattern of acquisition is marked and progressive from first through third and up to fifth year. This particular finding is highly relevant for my research, given that the experience of speakers in their L2 overcomes the social conditioning of the target form in Chilean Spanish, where the use of the voiceless alveolar fricative is found as a stigmatised realisation for voiceless alveolar affricate. Thus, the behaviour of this variable not only contradicts the hypercorrection hypothesis formulated initially, but it shows that the transfer of forms occurs only during the earlier stages of the acquisitional process. I argue that this is justified on the basis that speakers, after a considerable time of being exposed to English, are able to separate the two phonetic systems at the same time that they acknowledge the sociolinguistic value of the stigmatised feature of their L1 (otherwise, they would have kept on using the most prestigious variants over the less prestigious ones).

The case of voiced dental fricative is different: this is not as socially marked in the L1 of the speakers as the voiceless alveolar fricative, but in this it can be found as an allophone for /d/, predominantly in intervocalic and coda position, and within the

latter, it is linked to sociolinguistic conditioning (class, education; see Cepeda, 1991; Pérez, 2007). The results suggest that, despite being statistically significant, the intra-group variation is not homogenous; this means that inside groups the speakers exhibit dissimilar uses across the variants, therefore their individual behaviour cannot be considered characteristic of the whole group. In other words, if this is considered the factor that contributes the most to the variation of voiced dental fricative, it is because the difference between the mean of the three categories (first, third and fifth) represents the ideal scenario, but the actual behaviour of the sample shows that whereas some speakers have acquired the target form, others have not.

There is one interesting fact about the effect of this factor group on the variable: the alleged progression towards the acquisition of target forms seems to stop in third year, as third- and fifth-year speakers show similar rates of use. Recall Section 2.1.3.1.2 on interlanguage; Selinker (1972) suggests that there is a possibility that certain features remain fossilised as a result of the transfer from the L1 to the L2, forming part of the interlanguage or emerging system of the acquiring subject. This reference is highly applicable in this context, but the nature of the data prevents me from asserting so, thus, only future research will be able to tell whether the apparent stop on the progression of the target form for voiced dental fricative in connection to years of instruction in English results from the fossilisation of the feature in the interlanguage of the speaker.

Finally, for postvocalic (r) the first and most relevant finding relates to the use of target forms. The six variants found correspond to different realisations of the variable in different accents and dialects of English with different sociolinguistic values, but they can still be considered target form as they occur in native speech; the results, then, show that, the more time the speakers of the sample are exposed to formal instruction on the L2, the more they emulate target-like variation – but are they aware of this phenomenon? And is this correlated to the same (other) factors as in native speech? This, of course, depends on the accent they receive as input – in this case, the preferred variant corresponds to the canonical (characteristic) of GA (alveolar approximant), followed closely by another GA target form (retroflex approximant);[168] this connects to the second most relevant finding, which relates to

[168] Recall that the two majority variants correspond to the alveolar approximant, defined as canonical for GA, and the retroflex approximant, secondary GA target, as justified in Section 4.4.3.1.

the paradox devised from the use of the majority variant and the accent of instruction (refer to Section 4.4.2.4).

The overall results suggest that, firstly, the majority variant is not the canonical form of RP (in fact, this variant - zero realisation - accounted for only 6% of the total instances) but, rather, of GA (both, the alveolar and the retroflex approximants). Second, the expected progression pattern (first > third > fifth) does not occur: the use of the majority variant is preferred by fifth years, followed by third years. I propose as explanation a sort of u-shaped curve of rhoticity: in the earlier stages of acquisition of a new system, learners gradually incorporate features from the L2,[169] retaining the rhotic features of their L1 or assimilating them from GA. During intermediate stages of acquisition the internalisation of the L2 system (resulting from the increased exposure to the second language) is higher, therefore the learners may be more aware of the prestige of non-rhoticity, which explains the increase of the RP variant; however, the use of the two majority variants (both rhotic) increases in fifth-year learners to similar levels exhibited by first-year speakers. If progression has been tested for other variables, in this case the use of the rhotic variant does not mean that the learners show a step back in the acquisition process by using their L1 system, but that they are simply using features from a rhotic variety of English, in this case, GA.

Having tested the primary hypotheses, I now discuss the results for each variable, for the other conditioning factors and, specifically, for the hypotheses formulated for each of the four studied variables. I will focus on those factors that have not been discussed in the present section.

5.3 Testing secondary hypotheses: discussion of results per variable examined

The contents of this section are based on the comparison of the outcomes of previous research and the results drawn from the analysis of my data. As the first two variables share a common background in relation to the variation in English, I merge them together. The results for the remaining two variables are discussed individually.

[169] This assumption is made from the pedagogical point of view, but sustained from linguistics by Weinreich (1953), Flege (1995) and even interlanguage theory (Selinker, 1972).

5.3.1 Voiceless alveolar fricative (ʃ) and voiceless alveolar affricate (tʃ)

As shown in the literature review for (ʃ) and (tʃ), the variation of these two features has been studied primarily in native speakers of Spanish (see Sections 4.1.2 and 4.2.2, respectively). For native speakers of Spanish, voiceless alveolar fricative (ʃ) and voiceless alveolar affricate (ʧ) have been studied in Panama (Cedergren, 1973) and Chilean Spanish (Valencia, 1993; Valdivieso, 1983; Tassara, 1992; Cepeda, 2001), the first as a variant for the second, which has been found to be highly constrained by social factors, especially educational background and socioeconomic class. Sections 5.3.1.1 and 5.3.1.2 that follow, review and discuss the results drawn from the analysis of these two variables.

5.3.1.1 Social factors

For voiceless alveolar fricative (ʃ), given that one of its variant forms [ʃ] is a stigmatised feature for (ʧ), it was expected that the target form would be avoided, especially by lower-social-class learners. This was, to some extent, refuted: most speakers presented considerably high rates of use of the target form [ʃ], which suggests that they are aware that the two separate languages work independently from the social conditioning of each language. In other words, the informants do not seem to transfer the sociolinguistic value of (ʃ) from Chilean Spanish to English. This general pattern, however, presents a diversion that could be connected to social class: some lower-middle-class speakers disfavour the use of the target form (the stigmatised one) over the non-target affricate form, which could be interpreted as hypercorrection, as not only do they present higher rates of use of the non-target form, they do use the Chilean Spanish standard (which is prestigious) over the "fudge" form, which means that they identify the most prestigious form and use it, suggesting that this is not only a case of phonic interference. If it were the case of phonic interference, as suggested by Weinreich (1952), higher rates of the fudge form would have been found; this fudge form is a sort of intermediate solution to the realisation of two forms, as has also been argued by Cedergren (1973). However, intra-group variation was found to be highly heterogenous, therefore the high rates of use of the affricate variant by some speakers of the group does not represent the whole group variation. The heterogeneity of this group reveals that this factor does

not account for the source of variation in all speakers, but it is potentially an interesting group to test with a larger sample.

For voiceless alveolar affricate (ʧ): the basic premise here is of a different nature, as the target form (the voiceless alveolar affricate [ʧ]) is the standard for both English and Spanish; however, a new prestigious form occurs in Spanish and it is strongly present in my data (37%): an affricate form which emphasises the plosive over the fricative element [tʃ], being perceived more frequently as a voiceless alveolar plosive /t/.[170] The results suggest that this form, which has not been reported in English, has a considerable presence in the variation of (ʧ), confirming the transfer of this feature from Chilean Spanish to English.

This transfer seems to be justified from the socioeconomic factor group: the Chilean prestigious form shows higher rates of use in the lower-middle class, again possibly as a form of hypercorrection; they acknowledge the prestige of the fudge form in the L1 and extend its use to the L2. This factor group effect has also been enhanced by the interaction of years of education in English: the highest rates of use are in first-year speakers, which means that the presence and value of the L1 feature is stronger than in speakers from third and fifth year; thus, the effect of years of instruction in English, as laid out in the previous section, is confirmed.

5.3.1.2 Linguistic factors

With regards to the effects of linguistic constraints, the only linguistic factor group to exhibit relevant patterns of behaviour is the effect of the position of the variable in the word[171]. Recall that the effect of position, just as of preceding and following segment, seems to depend on the lexical item, therefore all results related to the articulatory effect of the variants will be subject to the characteristics of the word in which they occur. In the case of (ʃ), the occurrence of non-target forms was higher in initial position, which articulatorily favours strengthening, which here occurs by means of the affricate. The results for final position are also consistent with the literature: lenition and weakening have been widely reported to occur in less strong positions, which is especially justified for speakers from a Spanish L1 background, in

[170] It is recurrent in Chilean speakers of English being asked to repeat their place of origin after they are heard to say [ʻtɪli] instead of /ˈʧɪli/.

[171] These results are drawn from the analysis of voiceless alveolar fricative only.

whose system strong articulations – such as affricates and consonant clusters – rarely occur in the coda. However, is it possible that the articulatory phenomenon, in this case based on the L1 system of the individuals, is responsible for the variation? As shown above for the effect of the preceding and following segment, the variationist literature for both English and Spanish has found support in the study of articulatory phonetics to explain certain phenomena. In this case, based on articulatory phenomena such as lenition, it makes sense to find fricatives in stronger positions (initial) as fricative instances in middle and final positions. Although the results drawn from the analysis are not conclusive, they are worth noting for further extensions of the present research.

5.3.2 Voiced dental fricative

It has been shown that the variation of (ð) by native speakers of English is constrained by internal factors, such as preceding and following segment, prosodic position and lexical item,[172] and by external factors, such as class and ethnicity. In my data, only internal factors seem to contribute to the variation of (ð), particularly the position of the variable and the preceding segment.[173] This type of conditioning in NNSs then follows similar conditioning as for NSs. Now, specifically for native speakers of Spanish, the variation in the L1 is of a different nature, as this feature occurs as a variant for (d), frequently correlated to intervocalic and final positions. For position, I argued that, in articulatory terms, the occurrence of weaker forms in middle and final position in my data is consistent with the literature; this means that the use of non-target forms such as plosives (39%) in initial position versus the target fricative in middle and final position (98 and 81%, respectively) has, as in other studies and here with other variables, a relevant effect on the variation. This factor was statistically significant, however, only future research would be able to test if the correlation of position and the use of voiced dental fricative is relevant so as to account for the exhibited variation of this group of speakers.

5.3.3 Postvocalic (r)

[172] In Manuel (1995) and Baker (2008) the study of lexical item was related to word class distinction, which here could not be tested given the limited amount of lexical items that presented a considerable number of tokens.

[173] Sex, SEG and YIE tested significant, but intragroup behavior was considered non-homogeneous.

In native speakers of English, the variation of postvocalic (r) not only is the result of dialectal differences, but also of social and internal constraints. Of all the target forms,[174] three stand out: the zero realisation characteristic of RP (non-rhotic) and alveolar and retroflex approximants for GA (rhotic); the results showed that whereas the rhotic variants accounted for most of the tokens, the RP realisations occurred only minimally. Given the sociolinguistic behaviour of this feature in English, particularly its wide variation in the different accents, constrained by different factors, it cannot be contrasted on a one-to-one basis to non-native speech: in this context, the discussion of the variation of postvocalic (r) is focused on the effect of socioeconomic class – despite most speakers showing higher rates of rhoticity, it was lower-middle speakers who showed the highest rates of use of the RP variant. Then, how can these results be interpreted? A possible explanation lies in the greater access, therefore exposure, to the media by middle- and upper-middle-class speakers, who, unlike lower-middle-class informants, complement their formal instruction with (mostly) American TV shows and films; this, however, could not be tested with the present data, but Kuppens (2010) suggests that exposure to English-speaking TV shows (captioned in English) had an important effect on the acquisition of native-like forms for Flemish speakers, especially on girls (but not bound to class); however, her study is more focused on reading skills than on listening. The study of the effect of media on SLA variationist research is still incipient, but it is gaining space in native variation (see, e.g., Stuart-Smith, 2009; Lawson, Scobbie and Stuart-Smith, 2013).

5.4 Conclusions

Quantitative linguistics has provided the method and background for the examination of four phonetic variables in the speech of Chilean speakers of the English language; the results confirm that this variation is systematic and exhibits similar patterns of variation for the selected features in both English and Spanish, when applicable. The variation shown in the case of voiceless alveolar fricative and voiceless alveolar affricate cannot be contrasted to native-like variation, as in Durham's study,[175] as the occurrence of these features is restricted in some varieties of English as variants of

[174] The results showed six different variants for this variable, all of them corresponding to target forms, as they occur in different accents and dialects of English.

[175] Recall that, given the characteristics of the setting in which this research is conducted (non-naturalistic), it is considered a Type I variation study.

other variables. Furthermore, the linguistic behaviour of these Chilean Spanish speakers cannot be compared to speakers of other varieties of Spanish, as the language examined is English; and even between varieties, the sociolinguistic value of the examined features is of a different nature: what it is stigmatised in Chilean Spanish (voiceless alveolar fricative) is prestigious in Panama City. This is also valid for postvocalic (r), at least in what is concerned with the prestige of the variable: whereas rhoticity is stigmatised in some varieties of English, it is prestigious in others; therefore, non-native-like variation will only be comparable to native-like variation if this is restricted to one particular variety.

The case of voiced dental fricative is different: in both English and Spanish its internal conditioning shows similar patterning, therefore it is very difficult to determine if the patterns exhibited are as a result of acquired variation (Case A in Durham, 2006) or of variation influenced by the native language (Durham's Case B). Furthermore, would it be possible to assert, with the current data, that this group of individuals have acquired sociolinguistic competence, characteristic of Type II variation in a scenario that defines Type I variation (i.e., non-naturalistic)?

The distributional analysis has confirmed that the factors that contribute the most to the variation are: years of instruction in English, phonetic environment, and position. From the examination of the interaction of factors, the ones that tested to contribute to the variation are position in the word and preceding segments. Also, preliminary runs of MVA revealed that the effect of individuality for speakers and lexical item is significant for the variables examined.

5.5 Implications

Even though the effect of both social and linguistic constraints has been tested for voiceless alveolar fricative (ʃ), voiceless alveolar affricate (ʧ), voiced dental fricative (ð) and postvocalic (r), it is necessary to determine the possible implications of this study for variationist and second language acquisition research.

The results show that the variation in non-native speakers matches native variation in a number of different factors, especially in internal constraints; here, it seems that internal variation is acquired before external variation, which makes sense if we

consider that the context of acquisition is instructional and not of immersion. In other words, the acquisition of English in this group of students is not driven by the necessity to communicate that people face in immersion settings, but is focused on the acquisition of language as a system. The first implication of my study, then, is that if we, as speakers of EFL, acknowledge that variation is inherent to language, the paradigm of erroneous performance[176] will be overcome.

The second important implication deals with the informants as a group: they constitute the first group of Chilean speakers of English as a foreign language whose speech has been examined from the variationist research methodology perspective, under the premise to test that their patterns of sociolinguistic behaviour reflect to some extent their first language; potentially this study will set the basis to define the speech of Chilean speakers of EFL as one of the many varieties of English – such as Japanese English, Indian English, Chinese English – contributing with its own variants to the phonologic inventory of the word Englishes.

Finally, for instructional purposes, this study presents the notion that native-like features can occur in non-native speech, as tested in all variables studied, as a result of increasing exposure and experience to the L2. This is potentially important, as it constitutes a motivational resource for non-native speakers.

[176] It is pretty common to find the words 'erroneous' and 'mistake' in SLA research, when referred to the non-standard use of phonetic, morphosyntactic and discursive features.

REFERENCES

Abramson, A. S. and Lisker, L. (1972) Voice-timing perception in Spanish word-initial stops. *Haskins Laboratories Status Report on Speech Research*, *29*(30), pp. 15-25.

Adamson, H. D. (1989) Does the Bioprogram Affect Second-Language Acquisition?. In *The Dynamic Interlanguage* (pp. 33-47). Springer, US.

Adamson, H. D. and Kovac, C. (1981) Variation theory and second language acquisition. In D. Sankoff and H. Cedergren (eds.), *Variation omnibus*, pp. 215–25.

Adamson, H. D. and Vera M. Regan (1991) The Acquisition of Community Speech Norms by Asian Immigrants Learning English as a Second Language. *Studies in Second Language Acquisition*, 13, pp. 1-22.

ADIMARK (2000) El nivel socioeconomico Esomar. http://www.microweb.cl/idm/documentos/ESOMAR.pdf

Adjemian, C. (1976) On the nature of interlanguage systems. *Language Learning*, 26.2, pp. 297-320.

Al-Badawi, K. (2012) An Analysis of Phonetic, Morphological and Syntactic Errors in English: A Case Study of Saudi BA Students at King Khalid University. *International Journal of Social Science and Humanity*, 2.6

Alarcos-Llorach, E. (1958) Fonología y fonética (a propósito de las vocales andaluzas). *Archivum: Revista de la Facultad de Filología*, vol. 8, pp. 193-205.

Alvar, A. M. (1980). Los fonemas implosivos en español. *Thesavrvs,* Vol. 35, No. 3, pp. 456-505).

Anisman, P. H. (1975) Some aspects of code switching in New York Puerto Rican English. *Bilingual Review/La Revista Bilingüe*, pp. 56-85.

Anshen, F. S. (1969) *Speech Variation among Negroes in a Small Southern Community*. Doctoral dissertation, New York University, Graduate School of Arts and Science.

Anttila, A. (2002) Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* 20.1, pp. 1-42.

Babbitt, E. H. (1896) The English of the lower classes in New York City and vicinity. *Dialect Notes*, *1*(18), pp. 462.

Baker, R (2008) *The production and perception of morphologically and grammatically conditioned phonetic detail*. PhD Dissertation. University of Cambridge, Cambridge

Barboza, C.L.F. (2007) Production of English front vowels by Brazilian EFL teachers in Western Rio Grande do Norte. In Rauber, A. S., Watkins, M.A. and B.O. Baptista (eds.), *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speech. Florianópolis, Brazil: Federal University of Santa Catarina*, pp. 28-37.

Barron, A. (2003) *Acquisition in interlanguage pragmatics: Learning how to do things with words in a study abroad context* (Vol. 108). John Benjamins Publishing.

Bauer, L. and P. Warren (2004) New Zealand English: phonology. In Bernd Kortman and Edgar W. Schneider (eds.), *A Handbook of Varieties of English* Vol. 1. Berlin: Mouton de Gruyter, pp. 580-602.

Bayard, D. (1995a) Peers versus Parents: A Longitudinal Study of Rhotic and Non-rhotic Accents on a NZE Speaking Child. *New Zealand English Newsletter,* 9, pp. 15-22.

Bayley, R. (1994) Interlanguage variation in the quantitative paradigm: Past-Tense marking in Chinese-English. In: E. Tarone, S.M, Gass, and A. Cohen (eds.), *Research methodology in second language acquisition*. (pp. 157-181). Hillsdale, NJ: Lawrence Erlbaum.

_____. (2000) Second-language acquisition and variationist linguistics. .*American Speech*, *75*(3), pp. 288-290.

_____. (2002) The quantitative Paradigm. In: J.K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds.), *The Handbook of Language Variation and Change*. (pp. 117-141). Malden and Oxford: Blackwell Publishers.

_____. (2007) Second language acquisition: a variationist perspective. In Bayley R. And Lucas (eds.) *Sociolinguistic Variation, theories, methods and applications*. CUP.

Bayley, R. and Lucas, C. (eds.) (2007) *Sociolinguistic variation: theories, methods, and applications*. CUP.

Bayley, R. and Preston, D.R. (eds.) (1996) *Second Language Acquisition and Linguistic Variation*. Amsterdam and Philadelphia: Benjamins.

Bayley, G. and Thomas, E. (1998) Some aspects of African-American Vernacular English Phonology. In S.S. Mufwere, J. Rickford, J. Baught, and G. Bayley (eds.), *African-American English: Structure, History and use*. London: Routledge, pp. 85-109.

Bayley, G., T. Wikle, J. Tillery and L. Sand. (1991) The apparent time construct. *Language Variation and Change* 3, pp. 241-64.

Beebe, L.M. (1977) The influence of the listener on code-switching. *Language Learning* 27, pp. 331-339.

_____. (1980) Sociolinguistic variation and style shifting in second language acquisition. *Language Learning*, 30, pp. 433–445.

Bell, A. (1984) Language style as audience design. *Language in Society* 13, pp. 135–204.

Bell, A., and Johnson, G. (1997) Towards a sociolinguistics of style. *University of Pennsylvania Working Papers in Linguistics* 4, pp. 1-21.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., and Gildea, D. (2001) Form variation of English function words in conversation. *Submitted manuscript*.

Bell, T. S., and Wilson, R. H. (2001) Sentence recognition materials based on frequency of word use and lexical confusability. *Journal-American Academy of Audiology*, *12*(10), pp. 514-522.

Benson, B. (1988) Universal preference for the open syllables as an independent process in interlanguage phonology. *Language learning,* 38/2, pp. 221-35

Bernales, M. (1978) Sobre la palatal africada en el español de la ciudad de Valdivia. *RLA: Revista de lingüística teórica y aplicada*, (16), pp. 41-52.

Bernstein, J., (1993) *Topics in the syntax of nominal structure across Romance*. Ph.D. Dissertation, City University of New York.

Blondeau, Hélène, Nagy, Naomi, Sankoff, Gillian, and Thibault, Pierrette. (2002) La couleur locale du français L2 des anglo-Montréalais. In J.-M. Dewaele and R. Mougeon (eds.), *L'acquisition de la variation par les apprenants avancés du français langue seconde: Acquisition et Interaction en Langue Étrangère* 17, pp. 73–100.

Boberg, C. (2004) Ethnic patterns in the phonetics of Montreal English. *Journal of Sociolinguistics*, *8*(4), pp. 538-568.

Bonnici, L. M. (2010) *Variation in Maltese English: The Interplay of the Local and the Global in an Emerging Postcolonial Variety*. PhD dissertation, University of California. linguistics.ucdavis.edu/pics-and-pdfs/DissertationBonnici.pdf (10 October 2013)

Borland-Delorme, K. (2004) La variación y distribución alofónica en el habla culta de Santiago de Chile. *Onomazein* 10, pp. 103-115.

Boyce, S., and Espy-Wilson, C. Y. (1997) Coarticulatory stability in American English/r. *The Journal of the Acoustical Society of America*, *101*(6), pp. 3741-3753.

Butragueño, P.M. (2005) Listado de signos fonéticos, adaptación del AFI a la variación fónica del español. http://lef.colmex.mx/Sociolinguistica/Cursos/Listado%20de%20signos%20foneticos%202.pdf

Bybee, J. (2002) Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14(3), pp. 261–290

Campbell, R., and Wales, R. (1970) The study of language acquisition. *New horizons in linguistics*, pp. 242-260.

Canale, M., and Swain, M. (1980) Theoretical bases of com-municative approaches to second language teaching and testing. *Applied linguistics*, *1*(1), pp. 1-47.

Canfield, D. L., (1962) *La pronunciación del español de América*. Bogotá, Instituto Caro y Cuervo

Carlisle, R. S. (1997) The modification of onsets in a markedness relationship: Testing the interlanguage structural conformity hypothesis. *Language Learning*, 47, pp. 327-361.

Cartagena, N. (2002) Apuntes para la historia del español en Chile. In *Cuadernos de la Academia Chilena de la Lengua*. Santiago: Rumbos

Catford, J. C. (1977) *Fundamental problems in phonetics*. Midland Books.

_____. (2001) *A Practical Introduction to Phonetics* (Oxford Textbooks in Linguistics) 2nd (second) Edition. Oxford: OUP

Cedergren, H. (1973) *The interplay of social and linguistic factors in Panama*. Phd dissertation, Cornell University.

_____. (1987) The spread of language change: Verifying inferences of linguistic diffusion. In *Language spread and language policy: Issues, implications and case studies*. Washington, DC: Georgetown University Press, pp. 45-60.

Cedergren, H. and Sankoff, D. (1974) Variable rules performance as a statistical reflection of competence. *Language,* 50, pp. 233-255.

Cepeda, G. (2001) Estudio descriptivo del español de Valdivia, Chile, en *Estudios Filológicos* N.° 36, pp. 81-97.

Cepeda, G. and M. T. Poblete (1993) Retención y elisión de /β/ y /ð/ en sufijos y morfemas radicales, en *Estudios Filológicos* N.° 28, pp. 87-96.

Chambers, J.K. (1980) Linguistic variation and Chomsky's Homogenous speech community. *Papers from the fourth Annual Meeting of the Atlantic Provinces Linguistics Association, ed. Murray Kinloch and A. B. House*. Fredericton: University of New Brunswick, pp. 1-32.

_____. (2003) *Sociolinguistic Theory*. Malden, MA: Blackwell.

Chambers, J. K., and Trudgill, P. (1998) *Dialectology*. CUP.

Chand, V. (2010) Postvocalic (r) in urban Indian English. *English World-Wide*, *31*(1), pp. 1-39.

Cheshire, J. (1982) Linguistic variation and social function. *Sociolinguistic variation in speech communities*, pp. 153-166.

_____. (2002) Sex and gender in variationist research. In J. K. Chambers, P. Trudgill and N. Schilling-Estes (eds.), *The handbook of language variation and change*, 423–43. Oxford: Blackwell.

Cheshire, J., Fox, S., Kerswill, P., and Torgersen, E. (2008) Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistica*, *22*, pp. 1-23.

Cho, T. (2004) Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, *32*(2), pp. 141-176.

Chomsky, N. (1965) Aspects of the Theory of Syntax. *Cambridge, Mass*.

_____. (1993) *Lectures on government and binding: The Pisa lectures* (No. 9). Walter de Gruyter.

Contreras, L. (1988): Los anglicismos en el léxico del habla culta de Santiago de Chile, en ALFAL: VI (Phoenix, 1981), UNAM, 593-654.

Contreras, L., and Rabanales, A. (1991) Las muletillas en el habla culta de Santiago de Chile. In *Scripta philologica: in honorem Juan M. Lope Blanch* (pp. 673-744). Universidad Nacional Autónoma de México.

Contreras, M. (2002) Aspectos léxicos de la criollización lingüística, en *Estudios Filológicos* No 3 7. Valdivia: Facultad de Filosofia y humanidades, Universidad Austral de Chile, pp. 53-64.

Corder, S. P. (1967) The significance of learner's [sic] errors, *International Review of Applied Linguistics in Language Teaching* 5: pp. 161–70.

_____. (1981) *Error analysis and interlanguage* (Vol. 112). Oxford: OUP.

Cotton, E. G., and Sharp, J. M. (1988) *Spanish in the Americas*. Washington, DC: Georgetown University Press

Cruttenden, A. (1994) Rises in English. in J. Windsor-Lewis, ed. *Studies in general and English Phonetics: Essays in Honour of Professor J.D. O'Connor*. London: Routledge, pp. 155-173.

_____. (2001) *Gimson's pronunciation of English*. OUP.

Crystal, D. (1995) *The Cambridge encyclopaedia of the English language.* Cambridge: CUP.

Cucchiarini, C., and van den Heuvel, H. (1999) Postvocalic/r/-deletion in Dutch: more experimental evidence. In *Proceedings ICPhS San Francisco,* pp. 1673-1676.

Dalbor, J. B. (1959) The English phonemes /š/ and /č/: a hearing and pronunciation problem for speakers of Spanish learning English. *Language Learning*, 9, pp. 67–73.

Decamp, D. (1972) Hypercorrection and Rule Generalization. *Language in Society*, 1, pp. 87-90

Dewaele, J. M. (1992) L'omission du ne dans deux styles oraux d'interlangue française. *Interface. Journal of Applied Linguistics*, *7*(1), pp. 3-17.

_____. (2004a) Retention or omission of the ne in advanced French interlanguage: The variable effect of extralinguistic factors. *Journal of Sociolinguistics*, *8*(3), pp. 433-450.

_____. (2004b) The acquisition of sociolinguistic competence in French as a foreign language: An overview. *Journal of French Language Studies*, *14*(03), pp. 301-319.

Dickerson, L.J. (1975) The learner's interlanguage as a system of variable rules. *TESOL quarterly,* 9, pp. 401-7.

Dickerson, W. B. (1976) The psycholinguistic unity of language learning and language change. *Language Learning,* 26, pp. 215-31

Dickerson, L. and Dickerson, W. (1977) Interlanguage phonology: Current research and future directions. In S. P. Corder and E. Roulet (eds.), *The notions of simplification, interlanguages and pidgins*. Paris: Actes du 5 une colloque de Linguistique Appliqueede Neufchatel AIMAV/Didier, pp. 18-29.

Diéguez M., (2004) El anglicismo léxico en el discurso económico de divulgación científica del español de Chile, *Onomazein*, *10*(2), pp.117-141.

Doughty, C. J., and Long, M. H. (eds.). (2008) *The handbook of second language acquisition* (Vol. 27). John Wiley and Sons.

Dubois, S. and Horvath, B. (1998a) Let's tink about dat: Interdental fricatives in Cajun English. *Language Variation and Change,* 10, pp. 246–261.

_____. (1998b) From accent to marker in Cajun English: A study of dialect formation in progress. *English World-Wide,* 19, pp. 161–188.

Durham, M. (2006) *English in Switzerland: Inherent variation in a non-native speech community*. Doctoral dissertation, University of Fribourg, Switzerland.

Eckert, P. (1988). Sound change and adolescent social structure. *Language in Society*, *17*, pp. 183-207.

_____. (2008) Variation and the indexical field. *Journal of sociolinguistics*, *12*(4), pp. 453-476.

Eckert P. and J. Rickford (2001) *Style and Sociolinguistic Variation*. Cambridge, U.K.: CUP.

Edwards, J. and Beckman M.E. (2008) Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in phonological development. *Language Learning and Development*. Apr 1;4(2), pp. 122-156.

Eklund, R. and A. Lindström. (1998) How to handle foreign sounds in Swedish text-to-speech conversion: Approaching the 'xenophone' problem. *Proc. International Conference on Spoken Language Processing*, Sydney, 30 November–5 December, 1998, vol. 7, pp. 2831–2834.

Ellis, N. C. (1997) Vocabulary acquisition: word structure, collocation, word- class, and meaning. In N. Schmitt and M. McCarthy (eds.), *Vocabulary: description, acquisition and pedagogy*, 122–39. CUP.

Ellis, R. (1985a) *Understanding second language acquisition*. Oxford: OUP.

_____. (1987) Interlanguage variability in narrative discourse: Style shifting in the use of past tense. *Studies in Second Language Acquisition* 9, pp. 1-20

Fasold, R. (1972) *Tense Marking in Black English*. Washington, DC: Center for Applied Linguistics

Feagin, C. (1979) *Variation and change in Alabama English*. Washington DC: Georgetown University Press.

_____. (1990) The Dynamics of a Sound Change in Southern States English: From R-less to R-full in *Three Generations, Development and Diversity: Linguistic Variation across Time and Space*, ed. by J. Edmondson *et al*., SIL. University of Texas, Arlington. pp. 129-146.

Field, F. W. (2002) *Linguistic borrowing in bilingual contexts* (Vol. 62). John Benjamins Publishing.

Fillmore, L. W. (1991) When learning a second language means losing the first. *Early childhood research quarterly*, *6*(3), pp. 323-346.

Flege, J.E. (1991) Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America* 89, pp. 395–411.

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

Flege, J. E., Bohn, O.-S. and Jang, S. (1997) The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics* 25, pp. 437–70.

Flege, J. E. and Hillenbrand, J. (1984) Limits on phonetic accuracy in foreign language speech production. *Journal of the Acoustical Society of America* 76(3), pp. 708–721.

Flege, J., Munro, M. and MacKay, I. (1996) Factors affecting the production of word initial consonants in a second language. In Bayley, R. and D. Preston (eds.), *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins. pp. 47-73.

Flege, J., Munro, M. and MacKay, I. (1995d) The effect of age of second language learning on the production of English consonants. *Speech Communication*, 16, pp. 1-26.

Fougeron, C. (2001) Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*(2), pp. 109-135.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America*, *101*(6), pp. 3728-3740.

Foulkes, P., and Docherty, G. J. (2000) Another chapter in the story of/r: 'labiodental' variants in British English. *Journal of sociolinguistics*, *4*(1), pp. 30-59.

Frank, 1948. *The speech of New York City*. Ann Arbor, MI: University of Michigan Dissertation.

Gamkrelidze, T. V. (1975) On the correlation of stops and fricatives in a phonological system. *Lingua*, *35*(3), pp. 231-261.

Gardner, R. C. (1983) Learning another language: A true social psychological experiment. *Journal of language and Social Psychology*, *2*(2-3-4), pp. 219-239.

Gass, S. M. (2013) *Second language acquisition: An introductory course*. Routledge.

Gass, S. and L. Selinker. (2008) *Second language acquisition: an introductory course*. Routledge.

Gatbonton, E. (1978) Patterned phonetic variability in second-language speech: A gradual diffusion model. *Canadian Modern Language Review*, 34, pp. 335-347.

Gibson, A., and Bell, A. (2010) Performing Pasifika English in New Zealand: The case of bro'Town. *English World-Wide*, *31*(3), pp. 231-251.

Gimson, A. C. (1988 [1991]) ed *English Pronouncing Dictionary*: Daniel Jones, 14th Edn. (1977) Reprinted 1988 with revisions and supplement by Susan Ramsaran. Cambridge: CUP, 1991.

Goad, H., L. White and J. Steele. (2003) Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? *Canadian Journal of Linguistics* 48, pp. 243-263.

Goedertier, W., Goddijn, S., and Martens, J.-P. (2000) Orthographic transcription of the Spoken Dutch Corpus, Proceedings of LREC-2000, *Athens*, Vol. 2, pp. 909-914.

Goldsmith, J. (1981) Subsegmentals in Spanish phonology: An autosegmental approach W. Cressey, D. Napoli (eds.), *Linguistic symposium on Romance languages*, 9 Georgetown University Press, Washington, DC (1981), pp. 1–16

Goldstein, B. (2001) Transcription of Spanish and Spanish-influenced English. *Communication Disorders Quarterly*, *23*(1), pp. 54-60.

Gómez-Ortin, F. (2004) El dialect murciano y sus variedades. *Tonos Digital* (Revista electronica de estudios filológicos) Número 8 – Diciembre.

Gordon, M., and P. Ladefoged (2001) Phonation types: a cross-linguistic overview. *Journal of Phonetics*, 29, pp. 383-406.

Guenther, Frank H., Carol Y. Espy-Wilson, Suzanne E. Boyce, Melanie L. Matthies, Majid Zandipour and Joseph S. Perkell. 1999. Articulatory tradeoffs reduce acoustic variability during American English /r/ production. *Journal of the Acoustical Society of America* 105, pp. 2854–2865.

Guy, G.R. (1980) Variation in the group and the individual: The case of final stop deletion. In W. Labov (ed.), *Locating Language in Time and Space*. New York: Academic Press, pp. 1-36.

_____. (1988) Advanced VARBRUL analysis. *Linguistic change and contact*, pp. 124-136.

_____. (1989) Language and social class. *Linguistics: The Cambridge Survey: Volume 4, Language: The Socio-Cultural Context*, 4, 37.

_____. (1991) Explanation in variable phonology. An exponential model of morphological constraints. *Language Variation and Change* 3, pp. 1-22

_____. (1993) The quantitative analysis of linguistic variation. In D. R. Preston (ed.), *American Dialect Research*. Amsterdam: Benjamins.

_____. (1994) *Variable is variable: Principles, constraints, and linguistic variation.* Paper presented at NWAVE-XXIII, Stanford University,

_____. (2007) Lexical Exceptions in Variable Phonology, *University of Pennsylvania Working Papers in Linguistics*: Vol. 13: Iss. 2, Article 9.

Haensch, G (2001) Español de América y español de Europa (1a Parte). In *Panace@* Vol. 2, no. 6.

Hansen, J. (2001) Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics*, *22*(3), pp. 338–365.

_____.(2004) DEVELOPMENTAL SEQUENCES IN THE ACQUISITION OF ENGLISH L2 SYLLABLE CODAS: A Preliminary Study. *Studies in Second Language Acquisition*, *26*(01), pp. 85–124.

Hawkins, R., and Casillas, G. (2008) Explaining frequency of verb morphology in early L2 speech. *Lingua*, 118, pp. 595-612

Hernández-Campoy, J.M y J.M. Jiménez-Cano (2004) Análisis del proceso de estandarización lingüística en Murcia: el uso de archivos sonoros radiofónicos para su medición diacrónica y soncrónica. *Tonos Digital* (Revista electronica de estudios filológicos) Número 8 – Diciembre.

Hiang, T. C., and Gupta, A. F. (1992) Postvocalic /r/ in Singapore English. *York Papers in Linguistics* 16, pp. 139-152

Hickey, R. (1999) Dublin English: Current changes and their motivation. *Urban Voices: Accent Studies in the British Isles*, pp. 265-281.

Hindle, D. (1980) *The social and structural conditioning of phonetic variation*. University of Pennsylvania Ph.D. Dissertation.

Holmes, J. (1997) Women, Language and Identity. *Journal of Sociolinguistics*, 1, pp. 195–223.

Holmes, J., and Meyerhoff, M. (eds.). (2008) *The handbook of language and gender* (Vol. 25). John Wiley and Sons.

Horvath, B. (1985) *Variation in Australian English*. New York/Cambridge: CUP.

_____. (2004) Australian English: Phonology. In *A Handbook of Varieties of English: Vol. 1. Phonology*, edited by E. W. Schneider, K. Burridge, B. Kortmann, R. Mesthrie, and C. Upton. Berlin: Mouton de Gruyter, pp. 625-44.

Howard, M. (2004) Sociolinguistic variation and second language acquisition: a preliminary study of advanced learners of French *SKY Journal of Linguistics*, 17, pp. 134-165.

_____. Howard, M. (2010) S'approprier les normes sociolinguistiques en langue étrangere—quel role pour l'enseignement. *O. Bertrand*, *81*, pp. 283-96.

Howard, M., Mougeon, R., and Dewaele, J. M. (2013) *Sociolinguistics and second language acquisition*. Oxford Handbooks Online

Hubbell, A. F. (1950) *The pronunciation of English in New York City: Consonants and vowels*. New York: Columbia University's King's Crown Press.

Huebner, T. (1983) *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.

Huffman, M. K. (2005) Segmental and prosodic effects on coda glottalization. *Journal of Phonetics* 33, pp. 335–362.

Hughes, A, Trudgill, P. and Watt, D. (2005) *English Accents and Dialects: an Introduction to Social and Regional Varieties of English in the British Isles*, 4th edition. London: Hodder Arnold/New York: OUP.

Hughes, A. and Trudgill, P., (1979) *English Accents and Dialects*. London: Arnold

Hume, E. (2011) Markedness. *The Blackwell Companion to Phonology: Phonological interfaces*, *4*, 79.

Irwin, P., and Nagy, N. (2007) Bostonians/r/speaking: A quantitative look at (R) in Boston. *University of Pennsylvania Working Papers in Linguistics*, *13*(2), 11.

Jones, D. (1917) *An English Pronouncing Dictionary: (on Strictly Phonetic Principles)*. JM Dent and Sons, limited.

_____. (2011). *Cambridge English pronouncing dictionary*. P. Roach, J. Setter, and J. Esling (eds.). CUP.

Jongman, A., Wayland, R., and Wong, S. (2000) Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America*, 108, pp. 1252-1263.

Johnson, D. E. (2009) Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and linguistics compass*, *3*(1), pp. 359-383.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W.D. (2001) Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency and the emergence of linguistic structure,* ed. by Joan Bybee and Paul Hopper. Amsterdam: John Benjamins, pp. 229-254.

Kerswill, P., and Williams, A. (2000) Creating a new town koine: children and language change in Milton Keynes. *Language in society*, *29*(1), pp. 65-115.

Keune, K., M. Ernestus, R. van Hout and R. H. Baayen, 2005. Social, geo- graphical, and register variation in Dutch: From written 'mogelijk' to spoken 'mok'. *Corpus Linguistics and Linguistic Theory,* 1, pp. 183–223

Klein, W. (1986) *Second language acquisition*. CUP.

Kovecses, Z. (2000) *American English: an introduction*. Broadview Press.

Krashen, S. D. (1981) *Second language acquisition and second language learning* (Vol. 2). Oxford: Pergamon Press.

_____. (1988) Do we learn to read by reading? The relationship between free reading and reading ability. *Linguistics in context: Connecting observation and understanding*, 29, pp. 269-298.

Kuppens, A. H. (2010) Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, *35*(1), pp. 65-85.

Kutas, M., and Hillyard, S. A. (1983) Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition*, *11*(5), pp. 539-550.

Labov, W. (1963) The Social Motivation of a Sound Change. *Word* 19, pp. 273-309. Reprinted in *Sociolinguistic Patterns*. University of Pennsylvania Press, 1972, pp. 1-42.

_____. (1966a) *The Social Stratification of English in New York City.* Washington DC: Center for Applied Linguistics.

_____. (1966b) The linguistic variable as a structural unit. *Washington linguistics review* 3, pp. 4-22.

_____. (1969) Contraction, deletion and inherent variability in the English copula. *Language* 45, pp. 715-62.

_____. (1972a) *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

_____. (1972b) *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pensylvania Press.

_____. (1990) The intersection of sex and social class in the course of linguistic change. *Language Variation and Change*, 2, pp. 205-54.

_____. (1994) *Principles of Linguistic Change. Vol, I: Internal Factors*. Oxford and Cambridge, MA: Basil Blackwell.

_____. (2001) *Principles of Linguistic Change, Vol. II: Social Factors*. Oxford: Blackwell.

_____. (2010) *Principles of linguistic change, Vol. III: Cognitive and cultural factors*. Malden: Wiley-Blackwell.

_____. (2012) *The Sociolinguistic Intersection of Spanish and English.* James E. Alatis plenary talk at TESOL Convention, Philadelphia, March 29, 2012.

Labov, W., Ash, S. and Boberg, C. (2006) *The Atlas of North American English*. New York: Mouton de Gruyter.

Ladefoged, P. (1996) *Elements of acoustic phonetics*. University of Chicago Press.

_____. (2001) *Vowels and Consonants: An introduction to the sounds of languages.* Oxford: Blackwell Publishers.,pp. xxii+191.

Ladefoged, P., and Johnstone K., (2011) *A Course in Phonetics.* Cengage Learning (6th edition).

Ladefoged, P., and Maddieson, I. (1996) *The sounds of the world's languages*. Oxford: Blackwell.

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

Lavoie, L.M. (2002) Subphonemic and Suballophonic Consonant Variation: The Role of the Phoneme Inventory. *ZAS Papers in Linguistics*, 28, pp. 39 – 54

Lawson, E., Scobbie, J. M., and Stuart-Smith, J. (2011) A single case study of articulatory adaptation during acoustic mimicry. In: ICPhS XVII, 17-21 Aug 2011, Hong Kong http://eprints.gla.ac.uk/71771/

Levine, L., and H. Crockett Jr (1966) Speech variation in a Piedmont community: postvocalic r. In S. Lieberson (ed.), *Explorations in Sociolinguistics. Sociological Inquiry* 36, no. 2.

Levon, E. (2007). Sexuality in context: Variation and the sociolinguistic perception of identity. *Language in Society*, *36*(04), pp. 533-554.

Li, F., Edwards, J., and Beckman, M. E. (2009) Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers. *Journal of Phonetics*, 37(1), pp. 111-124.

Lipski, J. (1994) *Latin American Spanish.* New York: Longman.

Local, J. (1995) Phonological structure, parametric phonetic interpretation and natural-sounding synthesis. In *Fundamentals of speech synthesis and speech recognition.* John Wiley and Sons Ltd., pp. 253-270

Local, J. (2003) Variable domains and variable relevance: Interpreting phonetic exponents. *Journal of Phonetics*, 31, pp. 321-339

Macafee, C. (1983) *Varieties of English around the world: Glasgow*. Amsterdam: Benjamins

Macaulay, R.K.S (1977) *Language, Social Class, and Education: A Glasgow Study*. Edinburgh: The University Press.

MacLaughin, B. (1987) *Theories of second language learning*. London: Arnold.

Macwhinney, B. (2000. [2013]) *The CHILDES project. Tools for analysing talk*. Mahwah, NJ: Lawrence Erlbaum

Maddieson, I. (2011) Presence of uncommon consonants. In *The World Atlas of Language Structures Online* (ed. M. Haspelmath, M. Dryer, D. Gil, and B. Comrie), Chapter 19. Max Planck Digital Library, Munich.

Major, R. C. (1996) Markedness in second language acquisition of consonant clusters. In R. Bayley and Preston (eds.), *Second language acquisition and linguistic variation.* Benjamins Publishing Company, pp. 75-96

_____. (2004) Gender and stylistic variation in second language phonology. *Language Variation and Change*, 16, pp. 169-188.

Major, R. C. And Kim, E. (1996) The Similarity Differential Rate Hypothesis. *Language Learning,* 46, pp. 465–96.

Manuel, S. Y. (1995) Speakers nasalize/ð/after/n/, but listeners still hear/ð/. *Journal of Phonetics*, 23(4), pp. 453-476.

Manuel, S.Y., Shattuck-Hufnagel, S., Huffman, M., Stevens, K.N., Carlson, R., Hunnicut, S., (1992) Studies of vowel and consonant reduction. In: Ohala, J.J., Nearey, T.M., Derwing, B.L., Hodge, M.M., Wiebe, G.E. (eds.), *Proceedings of the 1992 International Conference on Spoken Language Processing,* pp. 943–946

Martínez, E., Planas, A. M. F., and Rivera, N. F. (2003) Pre-nuclear tonal inventories of Spanish intonation. In *Proceedings of the 15th International Congress of Phonetic Sciences*, pp. 595-598.

Meechan, M., and Foley, M. (1994) On resolving disagreement: Linguistic theory and variation–There's bridges. *Language variation and change*, *6*(01), pp. 63-85.

Melchers, G., and Shaw, P. (2003) *World Englishes: an introduction*. London, UK: Arnold.

Mendoza-Denton, N. (2007) *Home Girls*. Oxford, U.K.: Blackwell

Milroy, J. (1980) Lexical alternation and the history of English: evidence from an urbal vernacular. *Papers from the 4th International Conference on Historical Linguistics*, ed. Elizabeth C. Traugott, R. Labrum, and S. Sheperd. Amsterdam: John Benjamins, pp. 355-62.

_____. (1992) *Linguistic Variation and Chang*e. Blackwell.

_____. (2001) Language ideologies and the consequences of standardization. *Journal of sociolinguistics*, 5(4), pp. 530-555.

Milroy, J., and Milroy, L. (1978) Belfast: change and variation in an urban vernacular. In *Sociolinguistic Patterns in British English*, ed. Peter Trudgill. London: Edward Arnold, pp. 19-36.

Milroy, L., Milroy, J., and Docherty, G. (1999) Phonological variation and change in contemporary English: evidence from Newcastle upon Tyne and Derby. *Cuadernos de Filología Inglesa*, *8*.

Milroy, L. and Gordon, M. (eds.). (2003) *Sociolinguistics: method and interpretation.* Oxford: Blackwell.

Ministerio de Educacion. (2012) *Bases curriculares: Ingles como lengua extranjera*. http://www.curriculumenlineamineduc.cl/605/articles-21319_programa.pdf, accessed 10 October, 2013

Morosan, D. E., and Jamieson, D. G. (1989) Evaluation of a Technique for Training New Speech Contrasts: Generalization across Voices, but not Word-Position or Task. *Journal of Speech, Language and Hearing Research*, 32(3), pp. 501-511

Mougeon, R., and Rehner, K. (2001) Acquisition of sociolinguistic variants by French immersion students: The case of restrictive expressions, and more. *The Modern Language Journal*, *85*(3), pp. 398-415.

Mougeon, R., Nadasdi, T., and Rehner, K. (2010) *The sociolinguistic competence of immersion students* (Vol. 47). Multilingual Matters.

Mougeon, R., Rehner, K., and Nadasdi, T. (2004) The learning of spoken French variation by immersion students from Toronto, Canada. *Journal of Sociolinguistics*, *8*(3), pp. 408-432.

Myhill, J. (1988) Postvocalic/r/as an index of integration into the BEV speech community. *American Speech*, 63(3), pp. 203-213.

Nadasdi, T., Mougeon, R., and Rehner, K. (2005) Learning to speak everyday (Canadian) French. *Canadian Modern Language Review/La revue Canadienne des langues vivantes*, *61*(4), pp. 543-563.

Nagy, N., Blondeau, H., and Auger, J. (2003) Second language acquisition and real French: An investigation of subject doubling in the French of Montreal Anglophones. *Language variation and change*, *15*(01), pp. 73-103.

Nagy, N., and Irwin, P. (2010) Boston (r): Neighbo (r) s nea (r) and fa (r).*Language Variation and Change*, *22*(2), pp. 241-278

Noels, K. A., Pon, G., and Clément, R. (1996) Language, Identity, and Adjustment The Role of Linguistic Self-Confidence in the Acculturation Process. *Journal of Language and Social Psychology*, *15*(3), pp. 246-264.

Noriega, L. (2004) La fricativizacion de [š] en el español de Tucson, Arizona. *Divergencias - Revista de estudios linguisticos y literarios*, 2 (2), pp. 17-26

Odlin, T. (1989*) Language Transfer: Cross Linguistic Influence in Language Learning.* Cambridge, UK: CUP.

Ogden, R. (1996) Prosodies in Finnish. *York Papers in Linguistics* 17, pp. 191-240.

Oroz, R. (1966) *La lengua castellana en Chile*. Santiago: Editorial Universitaria.

Pérez, H. E. (2001) La noción de rasgo. El caso de las consonants oclusivas el espanol. *Onomázein*, pp. 327-336

_____. (2007) Estudio de la variación estilística de la serie/bdg/en posición intervocálica en el habla de los noticieros de la televisión chilena. *Estudios de fonética experimental*, 16, pp. 228-259.

Piercy, C. (2012) A Transatlantic Cross-Dialectal Comparison of Non-Prevocalic a/r/. *University of Pennsylvania Working Papers in Linguistics*, *18*(2), 10.

Pinker, S. (1994) *The language instinct: The new science of language and mind* (Vol. 7529). Penguin UK.

Piske, T., MacKay, I., and Flege, J. (2001) Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29, pp. 191-215

Poblete, M. T. (1995) El habla urbana de Valdivia: análisis sociolingüístico. *Estudios Filológicos.* Vol. 30. Valdivia, pp. 43-56.

Poplack, S. (1978) Dialect acquisition among Spanish-English bilinguals. *Language in Society* 7, pp. 89-103.

_____. (1980) Sometimes I'll start a sentence in English y termino en Español: toward a typology of code-switching. *Linguistics* 18, pp. 581-618.

_____. (1993) Variation theory and language contact. In *American dialect research: An anthology celebrating the 100th anniversary of the American Dialect Society*, pp. 251-286.

Poplack, S. and Tagliamonte, S. (2001) *African American English in the Diaspora*. Oxford: Basil Blackwell.

Poplack, S., Tagliamonte, S. and Eze, E. (2000) Reconstructing the source of Early African American English plural marking: a comparative study of English and Creole. In Poplack, S. (ed.) *The English history of African American English.* Oxford and Malden: Blackwell Publishers. pp. 73-105

Preston, D. (1985) The Li'l Abner syndrome. *American Speech* 60, 4, pp. 328-36.

_____. (1989) *Style, Status, Change: Three Sociolinguistic Axioms*. Paper given at NWAVE XVII.

_____. (1996) Variationist perspectives on second language acquisition. In R. Bayley and D. R. Preston (eds.), *Second language acquisition and linguistic variation.* Amsterdam: John Benjamins. pp. 1–45.

Preston, D. and Bayley, R. (2005 [1996]) *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins

_____. (2009) Variationist Linguistics and Second Language Acquisition. In W. Ritchie and T. Bhatia (eds), *The new handbook of second language acquisition*. Oxford: Elsevier.

Preston D. and Young, R. (2000) *Adquisición de segundas lenguas: Variación y contexto social.* (Ed. and translated by Francisco Moreno). Madrid: Arco Libros.

Rabanales, A. (2000) El español de Chile: presente y futuro, *Onomázein*, 5, pp. 135-151

Rabanales, A., and Contreras, L. (eds.). (1990) *El habla culta de Santiago de Chile: Materiales para su estudio* (Vol. 2). Instituto Caro y Cuervo.

Rauber, A. S., Escudero, P., Bion, R. A. H., and Baptista, B. O. (2005) *The interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese.* Paper presented at the EUROSPEECH-INTERSPEECH 2005 9th European Conference on Speech Communication and Technology, Lisbon.

Raymond, W. D., Dautricourt, R. and Hume, E. (2006): Word-Internal /t,d/ Deletion in Spontaneous Speech: Modeling the Effects of Extra-Linguistic, Lexical, and Phonological Factors, in: *Language Variation and Change* 18, pp. 55-97.

Regan, V. (1996) Variation in French interlanguage: A longitudinal study of sociolinguistic competence. In Robert Bayley and Dennis R. Preston (eds.), *Second language acquisition and linguistic variation* (pp. 177-201). Amsterdam: John Benjamins.

Rehner, K. (2010) The use/non-use of *ne* in the spoken French of university-level learners of French as a second language in the Canadian context. *Journal of French Language Studies*, *20*(03), pp. 289-311.

_____. (2011) The sociolinguistic competence of former immersion students at the post-secondary level: the case of lexical variation. *International Journal of Bilingual Education and Bilingualism*, *14*(3), pp. 243-259.

Rehner, K., Mougeon, R., and Nadasdi, R. (2003) The learning of sociolinguistic variation by advanced FSL learners. *Studies in Second Language Acquisition*, 25.1, pp. 127-156.

Riney, T. J., Takada, M. and Ota, M. (2000) Segmentals and Global Foreign Accent: The Japanese Flap in *EFL. TESOL Quarterly*, 34, pp. 711–737

Roach, P. (2004) British English: Received Pronunciation, *Journal of the International Phonetic Association* 34 (2), pp. 239–245

Rojczyk, A. (2010) Forming new vowel categories in second language speech: The case of Polish learners' production of English /i/ and /e/. *Research in Language* 8, pp. 85-97

Romaine, S. (1978) Postvocalic/r. *Scottish English: Sound change in progress*, pp. 144-157.

Romaine, S. (2008) Variation, in *The Handbook of Second Language Acquisition* (eds. CJ Doughty and MH Long), Blackwell Publishing Ltd, Oxford, UK

Sadowsky, S. and G. Salamanca. (2011) The phonetic inventory of Chilean Spanish: Guiding principles, provisional inventory of consonants and system of representation (IPA-CL). *Onomázein* 24 (2), pp. 61-84. (In Spanish)

Sáez, L. (1999) *El español de Chile en las postrimerías del siglo XX*. Santiago: Universidad de Santiago.

_____. (2005) Anglicismos en el español de Chile. *Revista Atenea n° 492, ii Sem, 2005*, Universidad de Concepción, Concepción, pp. 171-177.

Salamanca, G. and Marlet, S. (2001) *Curso básico de fonética general*. www.sil.org/training/capacitar/FONETICA/cursos/CursoBasicoFonetica.stm

Sankoff, G. (1974) A quantitative paradigm for the study of communicative competence. In *Explorations in the Ethnography of Speaking*, ed. R. Bauman and Joel Sherzer. CUP, pp. 18-49. Reprinted in *The Social Life of Language*, ed. Gillian Sankoff. University of Pennsylvania Press. 1980, pp. 47-80.

Sankoff, G., Thibault, P., Nagy, N., Blondeau, H. F., and MO, G. L. (1997) Variation in the use of discourse particles in a language contact situation. *Language Variation and Change*, *9* (2), pp. 191-217.

Sankoff, D. (1988b) Variable rules. In Ulrich Ammon, Norbert Dittmar, and Klaus J. Mattheier, eds., *Sociolinguistics: An international handbook of the science of language and society, vol. 2* (pp. 984-997)

Sankoff, D., Tagliamonte, S., and Smith, E. (2005) *Goldvarb X: A variable rule application for Macintosh and Windows.* Department of Linguistics, University of Toronto.

Santa Ana, O. (1996) Sonority and syllable structure in Chicano English. *Language Variation and Change* 8, pp. 3- 91.

Sapir, E. (1921) *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace and Company.

Sauer, W., (2002) S*ome recent developments in standard british english pronunciation and the teaching of EFL.* http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.413.3394andrep=rep1andtype=pdf

Schilling-Estes, N. (2001) *Variation study and discourse analysis: The sociolinguistic interview as a situated speech event*. Invited panel presentation, New Ways of Analyzing Variation (NWAV) 30, North Carolina State University, October 2001.

_____. (2006) Dialect variation. *An introduction to language and linguistics*, pp. 311-341.

_____. (2007) Sociolinguistic fieldwork. In Bayley, R. and C. Lucas (eds.) *Sociolinguistic Variation: Theory, Methods and Application*. (pp. 165-189). Cambridge: CUP.

Schmidt, R. (1987) Sociolinguistic variation and language transfer in phonology. In G. Ioup and S. H. Weinberger (eds.), *Interlanguage phonology* (pp. 365-377). Rowley, MA: Newbury House Publishers

Schönweitz, T. (2001) Gender and postvocalic /r/ in the American South: a detailed socioregional analysis. *American Speech* 76(3), pp. 259–285.

Schuman, J. H. (1978a): The acculturation model for second language acquisition en Gingras, R. C. (ed.), *Second Language Acquisition and Foreign Language Teaching* (Arlington, V. A. Center for Applied Linguistics).

Selinker, L. (1972) Interlanguage. *International Review of Applied Linguistics in Language Teaching* 10, pp. 209–31.

Selinker, L., and Douglas, D. (1985) Wrestling with 'context' in interlanguage theory. *Applied Linguistics*, 6(2), pp. 190-204

Shadle, C., Proctor, M. I., and Iskarous, K. (2008) An MRI study of the effect of vowel context on English fricatives. *Journal of the Acoustical Society of America*, 123(5), 3735.

Shockey, L., (1973) *Phonetic and phonological properties of connected speech.* Doctoral dissertation, Ohio State University

_____. (1978) Perceptual test of a phonological rule. *Haskins Laboratories Quarterly Progress Report*, pp. 147-50.

Shuy, R.W., Wolfram, W., and Riley, W.K. (1968a) *A study of social dialects in Detroit*. Washington, Dc: Educational Resources Information Center.

Silverstein, M. (2003) Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, *23*(3), pp. 193-229.

Slavoff, G. and Johnson, J. (1995) The effects of age on the rate of learning a second language. *Studies in Second Language Acquisition*, 17, pp. 1-16.

Smith, J. (2001) "You ∅ na hear o' that kind o' things": Negative do in Buckie Scots. *English World Wide*, 21 (2). Pp. 231-259

_____. (2007) Methods in sociolonguistic research: syntax and morphology. In: Llamas, C., Mullany, L. and Stockwell, P. (eds.) *The Routledge Companion to Sociolinguistics*. Routledge, London, UK

Smith, J. and Durham, M. (2011) A tipping point in dialect obsolescence? Change across the generations in Lerwick, Shetland. *Journal of sociolinguistics*, 15 (2). pp. 197-225.

_____. (2012) Bidialectalism or dialect death? Explaining generational change in the Shetland Islands, Scotland. *American Speech*, 87 (1). pp. 57-88

Smith, J., Durham, M., and Fortune, L. (2007) "Mam, my trousers is fa'in doon!": Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change*, *19*(01), pp. 63-99.

Smith, J., Durham, M., and Richards, H. (2013) The social and linguistic in the acquisition of sociolinguistic norms: caregivers, children, and variation. *Linguistics*, *51*(2), pp. 285-324.

Smith, J. and Tagliamonte, S. (1998) 'We were all thegither … I think we was all thegither': was regularization in Buckie English. *World Englishes* 17, pp. 105-126.

Stahlke, H. F. (2006) Assimilation to/r/in English initial consonant clusters. *English Today*, *22*(01), pp. 57-58.

Starks, D. and Bayard, D. (2002) Individual variation in the acquisition of post-vocalic /r/: daycare and sibling order as potential variables*. American Speech* 77, pp. 184-194

Strevens, P. (1985) Standards and the Standard Language. *English Today* 2(1(2)), 5

Stuart-Smith, J. (1999) Glasgow: Accent and voice quality. In P. Foulkes and G. Docherty (eds.), *Urban Voices*. London: Arnold, pp. 203-22.

_____. (2007) The influence of the media, in: Llamas, C., Mullany, L. And Stockwell, P. (eds.) *The Routledge Companion to Sociolinguistics. Routledge*, London, UK, pp. 140-148

_____. (2007) A sociophonetic investigation of postvocalic /r/ in Glaswegian adolescents. In: Trouvain, J. and Barrt, W. J. (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences: ICPhS XVI, 6 - 10 August 2007, Saarbrücken, Germany*. Universität des Saarlandes, Saarbrücken, Germany, p. 1307.

Stuart-Smith, J. and Timmins, C. (2010) The role of the individual in language variation and change. In In: Llamas, C. and Watt, D. (eds.) *Language and Identities.* Edinburgh University Press, Edinburgh, UK, pp. 39-54.

Stuart-Smith, J., Timmins, C. and Tweedie, F. (2007) 'Talkin' Jockney'?: variation and change in Glaswegian accent. *Journal of sociolinguistics*, 11 (2). Pp. 221-260.

Stuart-Smith, J., Timmins, C. and Wrench, A. (2003) Sex and gender in /s/ in Glasgwegian. In: Solé, M.J., Recasens, D. and Romero, J. (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences: 3-9 August 2003, Barcelona, Spain*. Causal Productions, Rundle Mall, Australia, pp. 1851-1854.

Tagliamonte, S. (1998) Was/were across the generations: view from the city of York. *Language variation and change*, 10 (2), pp. 153-92.

_____. (2006) *Analysing sociolinguistic variation*. Cambridge: CUP.

_____. (2007) Quantitative analysis. In Robert Bayley and Ceil Lucas (eds.), *Sociolinguistic variation: theories, methods and applications*. pp. 190-214. Cambridge: CUP.

Tagliamonte, S. A. and S. Molfenter (2007) How'd you get that accent?: Acquiring a second dialect of the same language. *Language in Society*, 36, pp. 649-675.

Tagliamonte, S., and Smith, J. (2005) No momentary fancy! The zero 'complementizer' in English dialects*. English Language and Linguistics*, 9 (2). Pp. 289-309

Tagliamonte, S. and Temple, R. (2005) New perspectives on an ol' variable: (t,d) in British English. *Language Variation and Change*, 17, pp. 281-302

Tarone, E. (1979) Interlanguage as chameleon. *Language Learning*, 29 (1), pp. 181-91.

_____. (1983) On the variability of interlanguage systems. *Applied linguistics*, 4, pp. 143-163.

_____. (1985) Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language learning* 35, pp. 373-404.

Tassara, G. (1992) Actitudes lingüísticas ante la variación de/ch. *Revista de Lingüística Teórica y aplicada*, *30*, pp. 263-271.

Thompson, R. (1976) Mexican-American English: Social correlates of regional pronunciation. *American Speech* 50, pp. 18-24

Timmins, C., Tweedie, F., and Stuart-Smith, J. (2004) Accent change in Glaswegian (1997 corpus): Results for consonant variables. *Department of English Language, University of Glasgow.*

Tomasello, M. (1992) *First verbs: A case study of early grammatical development*. CUP.

Trudgill, P. (1974) *The social differentiation of English in Norwich*. Cambridge: CUP.

_____. (ed.) (1985) *Language in the British Isles*, Cambridge: CUP.

_____. (1986) *Dialects in Contact*. Oxford: Blackwell.

_____. (1988a) Norwich revisited: recent linguistic changes in an English urban dialect. *English World-Wide* 9, pp. 33-49.

_____. (1990) *The dialects of English*. Oxford: Blackwell.

_____. (1992) *Introducing language and society*. London: Penguin.

Upton, C. (2008) 'Received Pronunciation', in Bernd Kortmann and Clive Upton (eds), *Varieties of English 1: The British Isles*. Berlin : Mouton de Gruyter, pp. 269-282

Valdivieso, H. (1981) Valoración subjetiva de usos lingüísticos, en *Actas del VI Congreso de ALFAL*, Phoenix, Arizona (EE.UU).

_____. (1983) Prestigio y estigmatización. Factor determinante en la ensenanza institucionalizada de la lengua materna. *RLA* 21.

_____. (1993) Perfil fonético de escolares de Concepción. *RLA* 31

Valencia, A. (1993-1994) Realizaciones de Išl, [č]/ y [ĵl en el habla adolescente. *Nueva Revista del Pacifico*, 38/39, pp. 159-180.

Van den Doel, R. (2010) Native and non-native models in ELT: Advantages, disadvantages, and the implications of accent parallelism. *Poznań Studies in Contemporary Linguistics*, *46*(3), pp. 349-365.

Van Petten, C., and Kutas, M. (1991) Influences of semantic and syntactic context on open-and closed-class words. *Memory and Cognition*, *19*(1), pp. 95-112.

Wagner, C. (1998) El atlas linguistico y etnografico de Chile por regions: ALECh. *Estudios filologicos*, No. 33, 1998, pp.119-130

_____. (1999) “Llevar a cuestas” en el atlas linguistico y etnografico de Chile (ALECh). *Estudios filologicos*, No. 34, 1999, pp. 193-200

_____. (2004) El atlas linguitico y etnografico de Chile: localidades y cuestionario. *Estudios filologicos*, No. 39, pp. 83-120

Wagner, S. E. (2008) *Linguistic change and stabilization in the transition from adolescence to adulthood* (Doctoral dissertation, University of Pennsylvania).

Weinreich, U. (1953) *Languages in Contact*. New York: Linguistic Circle of New York.

Weinreich, U., Labov, W. and Herzog M.I. (1968) Empirical foundations for a theory of language change. *Directions for Historical Linguistics: A symposium*, ed. Winfred P. Lehmann and Yakiv Malkiel. Austin: University of Texas Press, pp. 95-188.

Weiss, L. (1970) Auditory discrimination and pronunciation of French vowel phonemes. Dissertation *Abstracts international*. 31: 4051-A

Wells, J. C. (1982) *Accents of English 1: An introduction*. Cambridge, UK: CUP.

Wester, F., Gilbers, D., and Lowie, W. (2007) Substitution of dental fricatives in English by Dutch L2 speakers. *Language Sciences*, *29*(2), pp. 477-491.

White, L. (2003) Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition* 6, pp. 129-141.

Wigdorsky, L. (1978) Realización de algunos fonemas consonánticos en el castellano de Santiago: informe preliminar. *RLA* 16.

Wilde, L. F. (1995) *Analysis and synthesis of fricative consonants* (Doctoral dissertation, Massachusetts Institute of Technology).

Winford, D. (1978) Phonological hypercorrection in the process of decreolization–the case of Trinidadian English. *Journal of Linguistics*, *14*(2), pp. 277-291.

Wolfram, W. (1969) A Sociolinguistic Description of Detroit Negro Speech. Washington: *Center of Applied Linguistics*, 4, pp. 225-55.

_____. (1971) Black-White Speech Differences Revisited. In *Black-White Speech Relationships* ed. by Walter Wolfram and Nona H. Clarke, pp. 139-165.

_____. (1985) Variability in tense marking: a case for the obvious. *Language Learning* 35, pp. 229–53.

Wolfram, W., P. Carter and B. Moriello. 2004. Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics* 8, pp. 339-58

Wolfram, W., and Schilling-Estes, N. (2005) *American English: dialects and variation* (Vol. 20). Wiley-Blackwell.

Wolfson, N. (1976) Speech events and natural speech: Some implications for sociolinguistic methodology. *Language in Society*, 5, pp. 189-209.

You, H., Alwan, A., Kazemzadeh, A., and Narayanan, S. (2005) Pronunciation variations of Spanish-accented English spoken by young children. In *Interspeech*, 30, pp. 749-752.

Young, R. (1988) Variation and the interlanguage hypothesis. *Studies in second language Acquisition*, *10*(03), pp. 281-302.

_____. (1991) *Variation in interlanguage morphology*. New York: Peter Lang.

Young, Richard F. and Robert Bayley. (1996) VARBRUL analysis for second language acquisition research. In Bayley, R. and D. R. Preston (eds.), *Second Language Acquisition and Linguistic Variation* (pp, 253-306). Amsterdam: John Benjamins.

APPENDICES

Appendix A: Ethical approval sheet (sample)

Authorisation

Hereby, I authorise Mrs. Paulina Subiabre Ubilla to utilise the information obtained in the interview accomplished with academic purposes. Likewise, I authorised the utilisation of the aforementioned information to other members of the English Language Department of the University of Glasgow, for the same purposes, prior authorisation of Mrs. Paulina Subiabre.

I sign this, in complete knowledge that my identity will be kept in reserve, just as any other piece of information that could help to identify me.

Autorizacion

Mediante este documento, autorizo a la Sra. Paulina Subiabre Ubilla a utilizar la informacion obtenida en la entrevista realizada con fines academicos. De igual forma, autorizo la utilizacion de dicha informacion por terceras personas del departamento de Lengua Inglesa de la Universidad de Glasgow para iguales fines, previa autorizacion de la Sra. Paulina Subiabre Ubilla.

Firmo esta, en pleno conocimiento que mi identidad se mantendra en reserva, al igual que cualquier otro dato que me identifique.

Name / Nombre: _____

Interview Register / Numero de registro de entrevista:
_____

Signature / Firma: _____

Date / Fecha: _____

Appendix B: Advertisement

The advertisement referred to in the methodology section was posted in the hallways of the English language department of Universidad del Bio Bio; as it was written in English, it was addressed to a specific audience, this is, to students of English. As soon as the spaces were filled with the required number of speakers, the advertisements were removed. The advertisement read:

"Students are needed for an experiment in language use; your performance will not be assessed, is just talking!

"You will be asked to talk for about an hour to different questions on different topics. The only requirement is being a registered student in first, third or third year of the English teaching programme.

"If you wish to be part of it, email Paulina Subiabre on polisub@hotmail.com"

Appendix C: Interviewee data sheet (sample)

Interviewee Data

- Interview register: _____
- Name: _____
- Sex: _____
- Date of birth: _____
- Place of birth: _____
- Educational level:
- Elementary school: _____
- High school: _____
- College: _____
- Parents' educational level
- Mother: _____
- Father: _____
- Parents' occupation
- Mother: _____
- Father: _____
- Have you travelled abroad: yes ☐ no ☐
- Previous studies: _____

Appendix D: Questionnaire

This questionnaire intends to know about the motivation of Chilean undergraduate students towards English language acquisition. Some questions require a YES/NO answer, but it is expected an explanation for them or a full answer. The first part of the interview is related to the motivation behind the decision of learning English.

For first and third year students, the interview is centred in their experience with English language

For fifth year students, the initial questions are also repeated and the second part is related to their experience as high school teachers of English, and how they see the Chilean educational system regarding the teaching and learning of English as a foreign language.

FIRST & THIRD YEAR STUDENTS

- Why did you decide to study English?
- Have you ever thought of learning any other languages?
- When I left high school I studied one year of Architecture. Did you always think of studying English teaching?
- How did come you to study in Universidad del Bío-Bío?
- What's your home town like?
- In my hometown, people celebrate the anniversary by playing games and with big bonfires in the beach. Is there any special celebration in your home town?
- How do you usually spend holidays (national holidays, Christmas, New Year's Eve)? Which one is your favourite?
- Is there any special childhood memory?
- How do you find life as a student?
- How different is University from high school? (about responsibilities, friends, classmates and teachers)
- What was your favourite subject in high School?
- Most people write poems, short stories, songs, etc. when young, did you ever try writing?
- I was working in a tourist information office some years ago. Before entering the programme, did you face a situation in which you thought English was absolutely necessary to communicate with others?

- Have you ever travelled abroad?
- Have you ever been in an English speaking country? If so, how would you describe the experience?
- How would you describe your relationship with classmates and with teachers in university?
- What has been the most difficult part of learning English?
- What do you like/enjoy the most of learning English?
- Which of your classes is the favourite and why?
- Have you ever faced an embarrassing situation in class?
- How often and in which situations do you use English outside the classroom?
- Where and by what means do you receive most English input?
- (if the answer is related to written English) What kind of books/magazines/websites/ etc. you like the most? And Why?
- which is your favourite book or magazine? why
- (if the answer is related to spoken English) What kind of TV programmes, movies, music you prefer? Why?
- Which one is your favourite movie? Why?
- One if my favourite character is Elizabeth from Pride and Prejudice, both the book and the film. Do you have a favourite book or film character? Why is it your favourite?
- Have you notice any big difference between English and Spanish? (about structures, cultural differences)
- I have heard that the Chilean Government has stated for students of EFL at least one semester in an English speaking country, what you do think of that?
- Have you heard of the project "English Town"? (a fictitious town in the south of Chile in which public secondary students are intended to learn English in three months and they must speak only in English). What do you think of the idea?
- How do you see yourself in 5 more years?
- Would you like to teach in a high school or in a University? Why?
- Most people set short-term or long-term life objectives. Do you do it? Which is your biggest goal in life?
- Which is the most valuable experience you've ever had in life? And here, in the university?

FIFTH YEAR STUDENTS

- Why did you decide to study English?
- Have you ever thought of learning any other languages?
- When I left high school I studied one year of Architecture. Did you always thought of studying English teaching?
- How did come you to study in Universidad del Bío-Bío?
- What's your home town like?
- In my hometown, people celebrate the anniversary by playing games and with big bonfires in the beach. Is there any special celebration in your home town?
- How do you usually spend holidays (national holidays, Christmas, New Year's Eve)? Which one is your favourite?
- Is there any special childhood memory?
- How do you find life as a student?
- How different is University from high school? (about responsibilities, friends, classmates and teachers)
- I was working in a tourist information office some years ago. Before entering the programme, did you face a situation in which you thought English was absolutely necessary to communicate with others?
- Have you ever travelled abroad?
- Have you ever been in an English speaking country? If so, how would you describe the experience?
- How would you describe your relationship with classmates and with teachers in university?
- What has been the most difficult part of learning English?
- What do you like/enjoy the most of learning English?
- Which of your classes is the favourite and why?
- Have you ever faced an embarrassing situation in class?
- How often and in which situations do you use English outside the classroom?
- Where and by what means do you receive most English input?
- (if the answer is related to written English) What kind of books/magazines/websites/ etc. you like the most? And Why?
- Which is your favourite book or magazine? why

- (if the answer is related to spoken English) What kind of TV programmes, movies, music you prefer? Why?
- Have you notice any big difference between English and Spanish?
- I have heard that the Chilean Government has stated for students of EFL at least one semester in an English speaking country, what you do think of that?
- Have you heard of the project "English Town"? (Fictitious town in the south of Chile in which public secondary students are intended to learn English in three months and they must speak only in English). What do you think of the idea?
- My teaching practice was quite difficult. How did you find your teaching practice experience?
- While I was working in a public high school, English wasn't taken very seriously. How do you see the teaching of English in the Chilean educational system?
- Which do you think are the biggest flaws in the educational system concerning the teaching of English?
- Which do you think are the biggest obstacles in the teaching of English?
- Would you like to teach in a private o public high school (elementary school) or in an institute or university? Why?
- Most people set short-term or long-term life objectives. Do you do it? Which is your biggest goal in life?
- Which is the most valuable experience you've ever had in life? And here, in the university?

Appendix E: Coding Instructions

The following tables present the coding instructions for the four variables examined. Tables 1 (speaker, number of the interview register and code given), 2 (sex), 3 (years of instruction in English), and 4 (socioeconomic group) are common for the four variables.

Table 1. Speaker, number of interview register and code given

Speaker	Interview register number	Code
Paola	F1n1-03 PAB	3
Cecilia	F1n3-13 CRI	r
Marcela	F1n5-16 MAB	m
Consuelo	F3n4-10 KAR	0
Caterina	F3n2-14 KAT	k
Josefina	F3n6-17 JOS	j
Melinda	F5n1-11 MIL	i
Carola	F5n1-12 CAR	g
Carmen	F5n6-15 VAR	v
Bernardo	M3n1-1 VIC	1
Francisco	M1n1-2 FEL	2
Cesar	M5n1-4 CRI	4
Andres	M5n2-5 ADO	5
David	M3n2-6 DIE	6
Juan	M3n3-7 GON	7
Jaime	M1n2-8 JOS	8
Miguel	M5n3-9 GAB	9
Sebastian	M1n6-18 SAT	S

Table 2. Sex and codes given

Sex	Code
Female	F
Male	M

Table 3. Years of instruction in English of the speaker and codes given

Year	Code
First	1
Third	3
Fifth	5

Table 4. Socioeconomic group and codes given

Group	Code
C2 upper-middle	2
C3 middle-middle	3
D lower-middle	D

Coding instructions for Voiceless alveolar fricative (ʃ)

Table 5.1 Variants and codes given

Variant	Description	Code
ᵗʃ	Fricativised	m
tʃ	Stop	t
ʧ	Voiceless alveolar affricate	c
ʃ	Voiceless alveolar fricative	s

Table 5.2 Preceding phonetic environment, code given, and example

Preceding phonetic environment	Code	Example
(Long) pause/ full stop/ period	ç	JOSEFINA: “shy [s], i don't know there's a…”
Alveolar lateral approximant /l/	l	“greet people, shaking [s] hands, here in Chile…”
Alveolar tap /ɾ/	r	“sometimes scholarships [s] you have to pay a great…”
Back close /u/	u	“when you should [s] just read ahead…”
Back close-mid vowel /o/	o	“We wash [s] the dishes [s] and then…”
Bilabial nasal /m/	m	“and when I came she [s] said, her parents…”
Central close mid vowel /ɘ/	8	“you see a show [s] that is full, the tickets…”
Close front /i/ (short) & /iː/ (long)	i	“Everything is in English [sc], as should be…”
Front close-mid vowel /e/	e	“and very refreshed [s] magazines so I can learn”
Nasal Alveolar /n/	n	“(…) and this is one short [s] term objective…”
Near-close, near-back /ʊ/	0	“(…) her programme, so she [s] said, ok we are almost…”
Open front /a/	a	“I don't know, fashion [s], things like…”
Open mid-back vowel /ɑ/	h	“completely brushed [c] in all our bodies…”
Plosive Alveolar /t/	t	“I try to just short [s] objectives, reachable…”
Velar nasal (ŋ)	ñ	“was studying, she [s] couldn't come…”
Velar plosive /k/	k	“I like Shannia [c] Twain, and, James Blunt.”
Voiced alveolar plosive /d/	d	“from europe and she [s] knew spanish”
Voiced labiodental fricative /v/	v	“the kind of shows [s] that people like us…”
Voiceless alveolar fricative /s/	s	“that is shorter [s], but is one month I also…”
Voiceless labiodental fricative /f/	f	“reinvented herself, she [s] tried to make you…”

Table 5.3 Following phonetic environment, codes given, and example

Post phonetic environment	Code	Example
(Long) Pause/full stop/period	ç	"we talk in, in *Spanish* [t]."
Alveolar lateral approximant /l/	l	"language, *English [s] language*…"
Alveolar nasal /n/	n	"a very good *English [s] knowledge* but…"
Back close-mid vowel /o/	o	"is a kind of *show* [s] in the park there…"
Bilabial nasal /m/	m	"as a *punishment* [s] i mean"
Close front /i/ (short)	i	"said that *English [sc] is* the most important…"
Open-Middle central /ɜ/	3	"my t-*shirt [s*] got, i don't know how…"
Palatal approximant /j/	j	"but in *English [s] you* write in one way"
Voiced alveolar affricate /ʤ/	9	"thousand one I *finish [s] just* then the collection"
Voiced alveolar plosive /d/	d	"talking about *English [s], don't* we?"
Voiced bilabial plosive /b/	b	"I study English [s] because i like it"
Voiced dental fricative /ð/	6	"everything is in *English [s], the* environment is…"
Voiced labiodental fricative /v/	v	"speak in *Spanish [c] very* much"
Voiced labio-velar approximant /w/	w	"as I told you *English [s] where* i studied at the…"
Voiced velar plosive /g/	g	"learning to *English [s], grammatic* [sic] English"
Voiceless Alveolar Fricative /s/	s	"they live in *English [s] speaking* countries"
Voiceless alveolar plosive /t/	t	"…students of *Spanish [s] told* me."
Voiceless bilabial plosive /p/	p	"especially *British [s] people* are extremely…"
Voiceless labiodental fricative /f/	f	"my ideas in *English [s] for* me is so special."
Voiceless velar plosive /k/	k	"it makes *English [s] classes* hard to work."

Table 5.4 Lexical item and codes given

Lexical item	Code
British	b
English	&
Finish, finishing	f
Others (1 – 3 appearances)	o
Relationship	z
Share, sharing	y
She	q
Short, shorter, shorts	x
Should	a
Show, shown, shows	w
Spanish	$
T-shirt	t

Coding instructions voiceless alveolar affricate (ʧ)

Table 6.1 Variants and codes given for voiceless alveolar affricate

Variant	Description	Code
ᵗʃ	Fricativised	w
ʧˢ	Stop	x
ʧ	Voiceless alveolar affricate	ç
ʃ	Voiceless alveolar fricative	h

Table 6.2 Preceding phonetic environment, code given, and example

Preceding phonetic environment	Code	Example
Alveolar lateral approximant /l/	l	"(…) well or a *little shy* [s] when I'm nervous"
Alveolar tap /ɾ/	r	"I have a *scholarship [s*] so I don't need to"
Back close /u/	u	"Family. I like *to share* [s] with them"
Bilabial nasal /m/	m	"To try to *um share* [s] all the knowledge"
Central close mid vowel /ə/	8	"Um, once I wrote *a short [s]* story"
Close front /i/ (short) & /iː/ (long)	i	"She preferred S*panish [t]* Spanish, yeah."
Front close-mid vowel /e/	e	"And very *refreshed [s]* magazines so I can learn"
Nasal Alveolar /n/	n	"And but the r*elationship* [s] with my teachers"
Near-close, near-back /ʊ/	0	"Good, I *also shared* [s] with my classmates"
Open front /a/	a	"Er- and she *ha- she's* [s] not so child"
Pause	ç	*ADO: *"shorts [*s] stories, sayings…"
Plosive Alveolar /t/	t	*ADO*: "it should* [s] be included, it would be"
Velar nasal (ŋ)	ñ	ADO: "*everything should* [s] be taught in English"
Velar plosive /k/	k	"From Canada, so I *think she* [s] was very um good at teaching"
Voiced alveolar plosive /d/	d	"Good imagination *and she* [s] knows very well"
Voiced labiodental fricative /v/	v	"There is a sort *of show* [s] at in the main square"
Voiceless alveolar fricative /s/	s	"My childhood *was shown* [s], well, my character"
Voiceless labiodental fricative /f/	f	"She's *housewife, she [s*] doesn't work"

Table 6.3 Following phonetic environment, codes given, and example

Post phonetic environment	Code	Example
(Long) Pause/full stop/period	ç	"American *English* [s]." *PAU: well…
Alveolar lateral approximant /l/	l	"Eng-- *English [s] language* as a whole"
Alveolar nasal /n/	n	"get a very good *English [s] knowledge"*
Alveolar tap /ɾ/	r	"Teachers of *English [s], right*?"
Back close /u/	u	"Everything *should [*s] be taught in English"
Back close-mid vowel /o/	o	"Um, once I wrote a *short* [s] story"
Bilabial nasal /m/	m	"As a *punishment* [s] I mean"
Close front /i/ (short)	i	"(…) of- of knowing *English [sc] is* very important"
Front close-mid vowel /e/	e	"That's good, I also *shared* [s] with my classmates"
Open front /a/	a	"I didn't know *English [s] I* mean, I knew but what"
Open-Middle central /ɜ/	3	"Shorts and t-*shirts* [s], but, it's very beautiful"
Palatal approximant /j/	j	"The same, but in *English [s] you* write in one way"
Schwa	8	"All in *English [sc], and* we had another i me"
Voiced alveolar plosive /d/	d	"Talking about *English [s], don't* we?"
Voiced bilabial plosive /b/	b	"He speak only *English [sc], but* obviously when he…"
Voiced dental fricative /ð/	6	"History and *Spanish [s], those* were my favourite…"
Voiced labiodental fricative /v/	v	"Even speak in *Spanish [c] very* much, because…"
Voiced labio-velar approximant /w/	w	"Of American *English [s] was* stronger in me." *PAU
Voiceless Alveolar Fricative /s/	s	"And they live in *English [s] speaking* countries, so…"
Voiceless alveolar plosive /t/	t	"In secondary school, my *English [s] teacher* er, told me about it"
Voiceless bilabial plosive /p/	p	"Especially in the E*nglish [s] part*, because you, you…"
Voiceless glottal fricative/h/	x	"Subjects were *English [s], history*, and Spanish"
Voiceless labiodental fricative /f/	f	"Two student of *Spanish [s], foreign* students here"
Voiceless velar plosive /k/	k	"It could be an *English [s] colony*."

Table 6.4 Lexical item and codes given

Lexical item	Code
Chance	n
Chance	3
Change	2
Channel	j
Charge	k
Chat, chatting	i
Cheap, cheaper	6
Child, childhood	h
Children	c
Chile	*
Chilean	¡
Chillan	p
Chinese	d
Choice	4
Choose	5
Church	u
French	ñ
Much	m
Others (1 – 3 appearances)	o
Teach	e
Teacher	r
Teaching	g
Watch	s
Watching	S
Which	v

Coding instructions voiced dental fricative (ð)

Table 7.1 Variants and codes given for voiced dental fricative (ð)

Variants	Description	Code
[th1]	Laminal-interdental fricative	1
[th2]	Laminal-dentoalveolar fricative	2
[th3]	Apical-interdental fricative	3
[th4]	Apical-dentoalveolar fricative	4
[th5]	Laminal-interdental plosive	5
[th6]	Laminal-dentoalveolar plosive	6
[th7]	Apical-dentoalveolar plosive	7
[th9]	Voiceless laminal – interdental fricative	9
[th0]	deletion	0

Table 7.2 Preceding phonetic environment, codes given and examples

Preceding Phonetic Environment	Code	Example
(Long) pause/ full stop/ period	Ç	*Cri: "*they* [th1] answer with a smile, oh." F1N3-13 CRI
Affricate /ʧ/	6	"You just *approach them* [th6] and speak to them…"
All [a] vowels	A	"Yes, because, um, we, we *gather* [th1] all, the whole family it's, it's together…"
All [e] vowels	E	"We've been *together* [th1] for four years, and I spend a lot of time with him…"
All [i] vowels	I	"That is usually *with* [th1] mom, this year was different, I wasn't at home…"
All [o] vowels	O	"Well, I have four *brothers* [th1], we are five, counting me we are five, yes."
All [u] vowels	U	"*So there* [th1] were two or three days on the weekends…"
Alveolar nasal /n/	N	"*Even though* [th1] we broke, there're still the idea of meeting again so,…"
Alveolar tap	R	"That's kind of weird *for them* [th1], you know…" M5N1-04 CRI
Bilabial nasal /m/	M	"Oh, I have bad luck, because I' m finishing my *programme this* [th1] year."
Fricative /ʃ/	*	"Is very useful as I said before so, if I studied *English then* [th1] I can, I don't know study something else and I don't know."
Lateral /l/	L	"Maybe it's too late in high *school though* [th1]…"
Neutralised context (/t/)	0	"Mm, I think it's more easier for me *to write than* [th1] to speak."
Neutralised context (d)	̲º	"*And then* [th1] go abroad yes, and live there some years…"
Schwa /ə/	@	"So I *remember that* [th1] I was so sad, so sad and I had to, to go to one fifth year class and I got into"
Velar nasal /ŋ/	G	"Part of my family is *living there* [th2] and I go there and celebrate new year and Christmas, I hope"

Voiceless alveolar fricative /s/	S	"*Because this* [th1] institution has different areas, areas of education for children…"
Voiceless bilabial plosive /p/	P	"To *help them* [th1], I don't know, to get things they don't know how to get"
Voiceless labiodental fricative /f/	F	"So it's very important to travel to one *of those* [th1] countries."
Voiceless velar plosive /k/	K	"With my boyfriend or things *like that* [th1]."
labiodental fricative /v/	v	"In a private school when some *of them* [th1] have been in a native speaking…"

Table 7.3 Following phonetic environment codes, given and examples

Following Phonetic Environment	Code	Example
(Long) pause/ full stop/ period	Ç	"Professors, good, I haven't fought *with [th1*], but in with my classmates, ok" F3N2(5)-14 KAT
Affricate /ʤ/	Z	"Find a good, a good job and- and to feel good *with [th6] yourself*, to feel proud of yourself." M5N2-05 ADO
Affricate /ʧ/	6	"Is good to, to be *with [th6] children*, but is not so good being all the time with them," F5N1(5)-12 CAR
All [a] vowels	A	"*That's [th1]* why I know that…" F3N2(5)-14 KAT
All [e] vowels	E	"Ay, mm, I don't like *them* [th1], no." F1N2-03 . PAB
All [i] vowels	I	"In my school from, with all the school *within* [th1] Chillán." F3N2(5)-14 KAT
All [o] vowels	O	"Theatre in Santiago, and *agronomia* here in Chillan, but *those* [th1] were just for filling." F5N1(4)-11 MIL
Alveolar nasal /n/	N	"When I was in first year, at university i when to Mexico with [th4] nine friends…" M5N1-04 CRI
Alveolar tap	R	"[that's] the way I learn, you know *with [th1] rules* and everything…" M5N1-04 CRI
Bilabial nasal /m/	M	"Yep, but only to Argentina I went *with [th1] my* mother, the past winter" F3N6-17 JAS
Glottal fricative /h/	H	"I had to speak *with [th6] him*, but, but it wasn't totally necessary to know em, how can you say, a huge vocabulary" M3N2-06 DIE
Lateral /l/	L	"It's a wonderful place, um very old place with er- *with [th1] lots* of traditions, yes." M3N3-07 GON
Neutralised context (/t/)	0	"My relationship with, *with [th6] teachers* was good, in general…" M3NI-01 VIC
Neutralised context (d)	º	"Yes, er-, for, *with [th6] different* levels, for kids or in a bit a little older." F3N2(5)-14 KAT
Palatal approximant /j/	J	"It's funny, because you share *with [th2] your* friends, you can go out, you don't have so much responsibilities, yet" M1N1-02 FEL
Schwa /ə/	@	"They are nine *brothers* [th1] and sisters, so, it's ah, it's a big one." F3N2(5)-14 KAT
Voiced labial-velar fricative	W	"Er, yeah, *with [th1] words*, that I don't" M1N2(3)-08 JOS
Voiced velar plosive	G	"It's a drama *with [th1] good* things and bad things… "

/g/		M3N3-07 GON
Voiceless alveolar fricative /s/	S	"With teachers it's also good with, *with [th2] some* teachers, especially I have good relationships." M1N1-02 FEL
Voiceless bilabial plosive /p/	P	"When you spend time *with [th1] people* and you learn so fast and is more easy…" F1N3-13 CRI
Voiceless labiodental fricative /f/	F	"So imagine that, *with [th1] forty* five students." F5N6-15 VAR
Voiceless velar fricative /x/	X	"The bible. I have a bible *with [th5] have* the two idioms English and Spanish…" F1N5-16 MAB
Voiceless velar plosive /k/	K	"I don't have time to I don't know go parties with- *with [th1] classmates* for example." M3N3-07 GON

Table 7.4 Lexical item and codes given

Lexical Item	Code
Although	4
Another	A
Brother	B
Clothes	5
Clothing	C
Either	D
Father / grandfather	E
Further	F
Gather	G
Mother /grandmother	H
Neither	I
Other / otherwise	J
Rather	K
Rhythm	L
Than	M
That	N
The	Ñ
Their	2
Them	O
Then	3
There – therefore	P
These	Q
They	R
This	S
Those	T
Though	U
Together	V
Weather	W
Whether	X
With	Y
Within	Z
Without	1

Coding instructions postvocalic (r)

Table 8.1 Variants, description and codes given

Variants	Description	Code
[r1]	Alveolar flap – Spanish/Scottish "hará"/ "here"	1
[r2]	Alveolar approximant – GA standard "here"	2
[r3]	Zero – RP standard "here"	3
[r4]	Weakened alveolar approximant, almost zero	4
[r5]	Retroflex approximant	5
[r6]	Weakened alveolar flap	6

Table 8.2 Preceding phonetic environment, code given, and example

Preceding Phonetic Environment	Code	Example
All [a] vowels	A	"We don't have the attitude to go and talk, make mistakes, no, we *are* just too embarrassed, I would feel embarrassed…" (JAZ-F3)
All [e] vowels	E	"I like learning grammar, but now in third *year* we don't have grammar as a separate subject" (VIC-M3)
All [i] vowels	I	"I live *here* in Chillan, and I wanted to study er- near home" (VIC-M3)
All [o] vowels	O	"As I said *before* I want to teach people of this province" (GON-M3)
All [u] vowels	U	"A difficult situation a social one a lot of poor people student who do drugs…" (ADO-M5)
Schwa /ə/	@	"No, no, I feel *better* but I still don't feel confident" (SAT-M1)

Table 8.3 Following phonetic environment, code given, and example

Following Phonetic Environment	Code	Example
(Long) pause/ full stop/ period	,	Fel: "I have a little *sister*." (FEL-M1)
Affricate / ʧ/	C	"Universal history *or* Chilean history?" (PAB-F1)
Affricate /ʤ/	Y	"These holidays *are* just normal days" (VIC-M3)
All [a] vowels	A	"I wanted to be independent, *for* a while, and emancipate at least for four years" (JAZ-F3)
All [e] vowels	E	"The weather… it's hm, *or* extremely cold or extremely hot" (KAT-M3)
All [i] vowels	I	"With my classmates I try to use *more* English when we [are] out the class" (CRI-F1)
All [o] vowels	O	"You should be responsible for *your* own acts" (KAR-F3)
Alveolar nasal /n/	N	"The things as I say I *learned* in secondary school in arts are applied here" (GON-M3)
Approximant /x/	W	"Why *are* we going to speak English if we can speak

		Spanish" (CAR-F5)
Bilabial nasal /m/	M	"That was important *for* me, but now i love my university" (PAB-F1)
Fricative / ʃ/	Q	*"There* should be a lot more opportunities" (JAZ-F3)
Fricative /f/	F	"Because he had been *here* for two or three days" (CAR-M5)
Glottal fricative /h/	H	"You have to bring money to *your* house" (SAT-M1)
Lateral /l/	L	"I've always had a fascination *for* languages" (GAB-M5)
Palatal approximant /j/	J	"We could have something, but in five *more* years I think" (CAR-F5)
Plosive /d/	D	"They contain a lot *more* details" (JAZ-F3)
Plosive /p/	P	"Achieve my goals and to find a *better* position in a job" (ADO-M5)
Plosive /t/	T	"How do you think *your* teacher is" (CAR-F5)
Voiced bilabial plosive /b/	B	"I know if it's good *or* bad, but it's funny" (PAB-F1)
Voiced dentoalveolar fricative / ð/	6	"Was so good *for* this girl , but not for me" (CRI-F1)
Voiced interdental fricative /ɵ/	8	"She stayed *for* three months." (JOS-M1)
Voiced labiodental fricative /v/	V	"Because people *are* very irresponsible about their pets" (JAZ-F3)
Voiced velar plosive /g/	G	"For example after or before classes students *are* going to- to speak in Spanish" (GAB-M5)
Voiceless alveolar fricative /s/	S	"Could help me to open *doors* and things" (MAB-F1)
Voiceless velar plosive /k/	K	"They think in *work* in instead of continue studying" (CAR-F3)

Table 8.4 Lexical item and codes given

Lexical Item	Code assigned
(an)other	h
Are	a
First	b
For	c
Here	d
Learn(ing)	e
More	f
Or	g
Others	n
Suffixes	o
There	i
Understand	j
Work	k
Year	l
Your	m